Universidad Autónoma de Madrid

Departamento de Biología Molecular

Regulatory mechanisms of Germinal Centers

PhD Thesis

Arantxa Pérez García

Madrid, 2016

Regulatory mechanisms of Germinal Centers

Memoria presentada por la licenciada en Biología

Arantxa Pérez García

para optar al título de doctor por la Universidad Autónoma de Madrid

Directora de tesis:

Almudena R. Ramiro

Este trabajo ha sido realizado en el laboratorio de Biología de linfocitos B, en el Centro Nacional de Investigaciones Cardiovasculares (CNIC)

Madrid, 2016

Memoria presentada por **Arantxa Pérez García**, licenciada en Biología, para optar al grado de doctor por la Universidad Autónoma de Madrid.

Esta tesis ha sido realizada en el laboratorio de Biología de Linfocitos B del Centro Nacional de Investigaciones Cardiovasculares (CNIC), bajo la dirección de la **Doctora Almudena R. Ramiro**, y para que así conste y a los efectos oportunos, firma el siguiente certificado;

En Madrid, a 21 de Abril de 2016

Almudena R. Ramiro

RESUMEN

Tras el reconocimiento del antígeno, los linfocitos B pueden iniciar la reacción de centro germinal (GC), en la cual diversifican sus genes de inmunoglobulinas, mediante las reacciones de hipermutación somática (SHM) y cambio de isotipo (CSR), dando lugar a células plasmáticas o B memoria. La transición a través de los diferentes estadios de esta reacción implica la expresión coordinada de redes de genes que permiten una correcta diversificación de los linfocitos B. A nivel molecular, las reacciones de SHM y CSR se desencadenan por la desaminación de citosinas en los genes de las inmunoglobulinas, mediada por AID. La actividad de AID en linfocitos B no está restringida a los genes de las inmunoglobulinas, pudiendo introducir mutaciones en otros genes y mediar translocaciones cromosómicas con potencial linfomagénico. Además, AID puede expresarse en epitelio en respuesta a inflamación, sugiriendo un posible papel de AID en el desarrollo de carcinoma.

En la primera parte de este trabajo, se ha abordado el papel del remodelador de cromatina CTCF en la reacción de GC. Los datos obtenidos demuestran que CTCF es absolutamente necesario para el GC, pudiendo ser esta necesidad sobrepasada in vitro mediante la estimulación con citoquinas. El análisis transcripcional reveló que la deficiencia de CTCF desregula numerosos genes implicados en la reacción de GC así como en células plasmáticas, incluyendo Blimp-1. De hecho, el fenotipo asociado a la deficiencia de CTCF puede ser parcialmente revertido restableciendo los niveles de Blimp-1. En conjunto, los datos presentados sugieren que CTCF es un regulador maestro del GC, actuando en el mantenimiento del programa de GC y previniendo la diferenciación a célula plasmática.

En la segunda parte del trabajo, se ha testado el papel de AID en el desarrollo de carcinoma epitelial, mediante la generación de modelos murinos de sobre-expresión de AID en colon y páncreas. Mediante esta aproximación se vio que AID por sí misma no es capaz de promover la aparición de carcinomas, a pesar de introducir mutaciones y desencadenar una respuesta genotóxica. En cambio, AID promueve la expresión de ligandos NKG2D en páncreas, el reclutamiento de linfocitos $CD8^+$ y la inducción de muerte en el epitelio. Nuestros resultados indican que el potencial oncogénico de AID en células epiteliales puede ser neutralizado por un mecanismo protector de inmunovigilancia.

ABSTRACT

After the recognition of the antigen B lymphocytes can engage into the Germinal Center (GC) reaction, where they diversify their immunoglobulin (Ig) genes by somatic hypermutation (SHM), class switch recombination (CSR) and from where terminally differentiated plasma or memory B cells emerge. Transition through these different stages involves coordinated expression of gene networks that define transcription programs required for a proper B cell diversification. At molecular level, SHM and CSR are triggered by AID-mediated deamination of cytosines in Ig genes and can promote mutations and pro-lymphomagenic translocations. Importantly, AID activity in B cells is not restricted to Ig loci and can be expressed in response to inflammatory cues in epithelial cells, raising the possibility that AID mutagenic activity might drive carcinoma development.

In the first part of the work, the role of the chromatin remodeling factor CTCF during the GC reaction was addressed, as potential transcriptional regulator during this process. We found that CTCF is absolutely required for the GC reaction. Surprisingly, cytokine stimulation can rescue the need for CTCF for B cell activation in vitro. Transcriptome analysis has revealed that CTCF deficiency deregulates numerous genes involved in the GC reaction and in plasma cell, including Blimp-1. Indeed, we have found that CTCF deficiency is partially rescued by re-establishing Blimp-1 levels. Together, our data suggest that CTCF is a master transcriptional regulator of the GC reaction by maintaining the GC program and preventing plasma cell differentiation.

In the second part, the role of AID in carcinoma development was tested by generating conditional knock-in mouse models for AID overexpression in colon and pancreas epithelium. AID overexpression alone was not sufficient to promote epithelial cell neoplasia in these tissues, in spite of displaying mutagenic and genotoxic activity. Instead, we found that heterologous AID expression in pancreas promotes the expression of NKG2D ligands, the recruitment of $CD8^+$ T cells and the induction of epithelial cell death. Our results indicate that AID oncogenic potential in epithelial cells can be neutralized by immunosurveillance protective mechanisms.

INDEX

ABBREVIATIONS

Abbreviation	Full name
AID	Activation-induced deaminase
Bcl-6	B-cell lymphoma 6
BCR	B cell receptor
Blimp-1	B lymphocyte-induced maturation protein-1
CIITA	MHC class II transactivator
CSR	Class switch recombination
CTCF	CCCTC- binding factor
DSB	Double-strand break
GC	Germinal Center
H2AX	Histone 2A family-member X
hCD2	Human cluster of differentiation 2
Ig	Immunoglobulin
IL4	Interleukin-4
LPS	Lipopolisacharide
mRNA	Messenger RNA
NHEJ	Non-homologous end joining
NK	Natural killer cell
NKG2D	Natural killer group 2 member D
PC	Plasma cell
PP	Peyer's Patches
qRT-PCR	Quantitative real time polymerase chain reaction
RAE	Retinoic acid early inducible
SHM	Somatic Hypermutation
SRBC	Sheep red blood cells
SSB	Single-strand break
TCR	T cell receptor
Tfh	T follicular helper cells
Th	T helper cell
UNG	Uracyl DNA glycosilase
WB	Western Blot

INTRODUCTION

1. Immune system and adaptive immunity

The immune system is a very sophisticated mechanism to protect organisms against foreign agents. A key feature of the immune response is the identification of factors as potentially pathogenic. All animals possess a primitive system of defense called innate immunity, which is based in the recognition of molecular patterns commonly present in pathogenic agents. The innate response is a first rapid, efficient and nonspecific barrier against pathogens, and allows their efficient elimination in an unspecific way. In addition, it plays an important role in the activation of the adaptive immune response. This adaptive response is a most sophisticated system of defense that has evolved in vertebrates. The hallmark of the adaptive immune response is the diversity of their antigen receptors that allow a high specific recognition of pathogens and the development of immune memory, what makes possible a faster and more efficient elimination of the pathogen in future infections.

T and B lymphocytes, or cells, are the effector cells of the adaptive immunity. T cells are involved in the elimination of intra-cellular pathogens and in the activation of a number of immune cells. On the other hand, B cells are specialized in the development of humoral immune responses against extra-cellular pathogens. The function of both T and B cells relies on the expression of extremely diverse receptors for antigen T cell receptor (TCR) and B cell receptor (BCR), that are able to specifically recognize molecular structures (antigens) present in pathogens. Therefore each particular cell expresses a unique receptor and together the lymphocyte system provides an immense repertoire of different receptors able to detect and eliminate any potential infection.

2. B cell differentiation and antibody diversification

B cell differentiation allows the generation of B cells carrying a functional BCR. The BCR, or antibody in its secreted form, is a multiproteic complex comprised by a membrane-bounded immunoglobulin responsible for antigen recognition and a heterodimer Igα/Igβ that is involved in the signal transduction. Signaling through this receptor modulates a number of aspects of B cell biology.

Immunoglobulins are compound of two pairs of identical chains, called light chain (IgL) and heavy chain (IgH). The N-terminal region of both IgL and IgH constitute the variable region of the antibody, involved in antigen recognition. The C-terminal region of the IgH chain encodes the constant region of the antibody, which determines the elimination pathway of the antigen.

The variable region of the immunoglobulin is composed of numerous gene segments, which encode the variable (V), diversity (D) and, in the case of IgH, joining (J) exons. The constant region is formed by eight gene segments (Cμ, Cδ, Cγ1, Cγ2a, Cγ2b, Cγ3, Cε and Cα) which code for different constant regions.

2.1. Early B cell differentiation and primary antibody diversification

The genetic information encoded in the immunoglobulin genes is not enough to explain their high diversity. The diversification of immunoglobulins is achieved in two stages. During the process of B cell development in the bone marrow, there is a first mechanism of antibody diversification called V(D)J recombination. This antigen-independent process is a site specific recombination reaction, mediated by the recombinase activating gene (RAG) complex, between recombination signal sequences (RSS) that flank each of the V, D and J gene segments. RAG recombinases initiate the reaction by introducing a double strand DNA break in the RSS of each of the sequences to be joined. This break is repaired by the non-homologous end-joining (NHEJ) pathway, generating a gene unit composed by one of each

VDJ segments. During the VDJ recombination, the diversification in the immunoglobulin repertoire arises from the aleatory selection of the different VDJ segments and the inexact joining of the different regions during the repair mediated by the NHEJ pathway. This process generates a primary repertoire composed by 10^5-10^6 specificities (reviewed in (Bassing et al., 2002)).

2.2. Late B cell differentiation and secondary antibody diversification

B cells that have successfully rearranged their immunoglobulin genes by V(D)J recombination express IgM/IgD in their membrane and migrate to the periphery where the antigen-encounter triggers a secondary diversification of the antibody. This second step in antibody maturation occurs in secondary lymph organs in structures called germinal centers (GC) and is essential to generate high-affinity antibodies.

Secondary antibody diversification in GCs comprises two molecular reactions, called SHM and CSR. SHM introduces point mutations in the DNA encoding the variable region of the Ig. These mutations generate a collection of related B cell clones from which those with increased affinity for the cognate antigen are positively selected by affinity maturation (reviewed in (Di Noia and Neuberger, 2007)). CSR is a region-specific recombination that involves the replacement of the Cμ or Cδ (which code for the IgM or IgD heavy chains), primary expressed, by downstream constant regions Cγ, Cε or Cα (which code for the IgG, IgE or IgA heavy chains). CSR allows the combination of the same N-terminal region of the antibody with different C-terminal regions, thus diversifying the effector capabilities of Ig of a given antigen specificity (reviewed in (Xu et al., 2012)). Although molecularly different, both SHM and CSR are triggered by the action of the same enzyme, activation-induced deaminase (AID) (Muramatsu et al., 2000). Details about AID action will be discussed below.

Secondary diversification in GCs can give rise to 10^{12}-10^{13} BCR specificities and thus fundamental for the production of high-affinity antibodies and for a proper immune response.

3. Germinal Centers

GCs emerge in follicles present in secondary lymphoid organs, and involve intense cell proliferation and cell death, derived from the clonal expansion and affinity maturation processes. While absolutely required for a proper immune response, GCs are a prone environment to malignant transformation; indeed, the vast majority of lymphoid neoplasms originate from GC B cells (reviewed in (Kuppers, 2005)).

Secondary lymphoid organs are structured in follicles composed mostly by naïve B cells that are surrounded by T cells. After antigen encounter, naïve B cells migrate to the T rich zone of follicles in secondary lymph organs, where they are fully activated by its interaction with specificity-matching $CD4^+$ T cells. This interaction is mediated mainly by CD40, a member of the TNF receptor family expressed by B cells, with the CD40 ligand (gp39) expressed in T cells. B cells thus activated can give rise directly to short-lived antibody-secreting cells or enter the GC program.

The initiation of the GC reaction involves an intense process of cell proliferation that allows the clonal expansion of the cell that recognized the antigen. At the same time, B cells diversify their immunoglobulin genes by SHM (see below) and generate a pool of clones with different affinities. B cell clones will be selected based on their affinity for the initiating antigen by signals emanating from their surface Ig (BCR) in the context of Tfh and dendritic cells. Ultimately, further rounds of proliferation, SHM and selection will lead higher affinity variants to terminally differentiate into memory B cells or antibody secreting plasma cells (reviewed in (Victora and Nussenzweig, 2012)), The exit of the GC reaction involves a switch in the transcriptional program that halts cell cycle progression, as well as Ig diversification, and allows the differentiation to antibody secreting plasma cells (PC), the effector cells of the humoral response, and memory B cells (Figure 1).

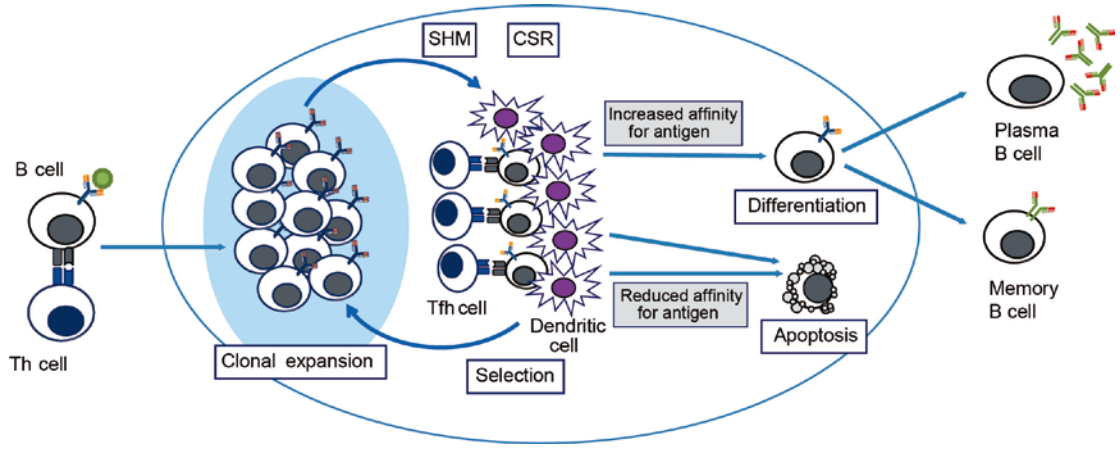

Figure 1. Germinal Center dynamics. In the periphery, naive B cells that have successfully undergone V(D)J recombination and express a functional B cell receptor can enter the GC reaction after antigen encountering. Naïve B cells become fully activated after interacting with Th cells. This activation triggers the aggregation and the clonal expansion of B cells, fan initiates the GC reaction. In the GC, B cells proliferate and remodelate their Ig genes by SHM and CSR and generate clones with higher affinity for the antigen that triggered the reaction by affinity maturation (see text for more details).

3.1. Transcriptional regulation of the Germinal Center

A number of key transcription factors have been associated with the initiation, maintenance and exit of the GC reaction. c-Myc is expressed in hyperproliferative populations of GC B cells both in immature and mature GCs and it is essential for GC formation and maintenance (Calado et al., 2012). When GCs are established, Bcl-6 is considered the master regulator of their transcription program maintenance. Indeed, Bcl-6 deficient mice do not generate GCs or high affinity antibodies and conversely, Bcl-6 overexpression increases GC formation. Bcl-6 is a transcriptional repressor that directly suppresses p53 expression, and through its binding to PIAS2, suppresses also the activation of the cell cycle-arrest gene p21. Thus, the transcriptional program modulated by Bcl-6 might allow the high proliferation rate present in GC B cells as well as tolerance to DNA

damage (reviewed in (Basso and Dalla-Favera, 2010)) (Figure 2). Other factors are involved in the maintenance of proliferation. EZH2 expression in mature B cells is restricted to GC B cells and is required for full GC development by promoting B cell proliferation and preventing their terminal differentiation, by inhibiting Blimp-1 (Caganova et al., 2013). In addition, E2A, a basic helix-loop-helix transcription factor, promotes cell cycle progression by inducing CCND3 and E2F2, and by downregulating RB1 (Schmitz et al., 2012).

The exit from the GC and the generation of antibody secreting PC involves a switch of the transcriptional program and requires the downregulation of Bcl-6, which acts repressing Blimp-1, a master regulator of the PC program. Blimp-1 expression is necessary to mediate plasma cell differentiation (Shapiro-Shelef et al., 2003) by silencing genes that specify B cell identity, such as surface B cell markers, B cell activation markers, B cell-associated transcription factors and genes induced in BCR signaling. It also represses CIITA, which leads to the downregulation of MHC class II genes and antigen presentation (Piskurich et al., 2000, Silacci et al., 1994, Shaffer et al., 2002) (Figure 2). In addition, Blimp-1 downregulates genes involved in cell proliferation, partly by directly repressing c-Myc.

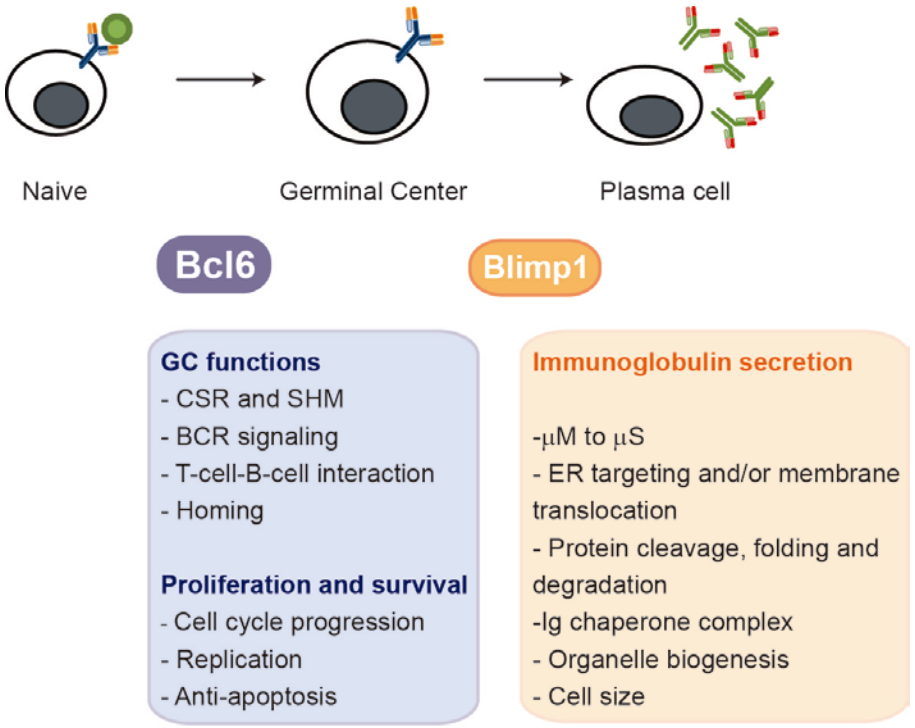

Figure 2. Germinal center and plasma cell transcriptional program. Master regulators of the Germinal center (Bcl6) and the Plasma cell (Blimp-1) programs are depicted. In the text boxes below, the main biological processes controlled by these factors are summarized. Among the pathways controlled by Bcl6, immunoglobulin remodeling through CSR and SHM, as well as cell cycle progression and replication can be highlighted. On the other hand, Blimp-1 is responsible of the differentiation to antibody secreting cells, favoring all the processes involved in protein synthesis and processing.

4. CCCTC-binding factor

CCCTC-binding factor (CTCF) is a highly conserved protein ubiquitously expressed in higher eukaryotes (Klenova et al., 1993). It contains a conserved 11 zinc finger DNA-binding domain by which it interacts with 50000-65000 sites in mammalian genomes (Chen et al., 2012). CTCF was initially described by Lobanenkov and colleagues as a DNA-binding protein involved in the regulation of the chicken c-myc gene (Lobanenkov et al., 1990). A later work carried by Filippova and colleagues defined CTCF as a multivalent factor by its flexibility in Zn finger usage in the interaction with the chicken and human c-myc promoter, and showed that it repressed c-myc transcription (Filippova et al., 1996). Initially, analysis of CTCF activity in a particular promoter using reporter transgenes suggested that CTCF acted as a barrier to transcription progression (Kellum and Schedl, 1991). Later on, CTCF was shown to block enhancer activity when placed between enhancers and promoters, indicating that CTCF binding sites could act as transcriptional insulators (Giles et al., 2010, Phillips and Corces, 2009). However, recent genome-studies suggest that CTCF could determine the establishment of functional expression domains, rather than act as a physical chromatin barrier (Handoko et al., 2011). The use of Chromosome Conformation Capture Carbon Copy (5C) has revealed the existence of long-range interactions that include enhancers, promoters and CTCF-binding sites. CTCF binding to these sites did not promote physically insulation of gene domains; rather, a proportion of these sites were enriched in active transcription marks (Sanyal et al., 2012). Several evidences support that interaction between promoters and enhancers is the main activity for CTCF. For example, transcriptional activation of the major histocompatibility complex II (MHC II) is preceded by a DNA looping event that allows the interaction between the XL9 enhancer and its cognate promoter. This looping has been shown to be mediated by CTCF, the MHC class II transactivator (CIITA) and specific transcription factors (Majumder and Boss, 2010, Majumder et al., 2008). CTCF has also been involved in the regulation of the transcription of the human protocadherin gene A (PDAC) cluster, which contains a number of exons that are stochastically expressed by alternative promoters. Promoter choice is mediated by CTCF-cohesin DNA looping, which forms and active chromatin hub. The binding of CTCF to the promoters preceding each exon correlated

with the level of gene activity (Golan-Mashiach et al., 2012, Kehayova et al., 2011, Guo et al., 2012, Hirayama et al., 2012, Monahan et al., 2012). All these evidences support the role of CTCF as a general transcriptional regulator, by mediating long-range interactions between promoters and distal enhancers.

4.1. CTCF in B cells

CTCF is a fundamental factor controlling the process of B cell differentiation in the bone marrow, mainly by controlling the process of V(D)J recombination. The first evidences of the role of CTCF in the regulation of V(D)J came from the identification of a regulatory region located between V and D clusters, called IGCR1, that underwent CTCF-mediated looping to bring together promoters and enhancers. This looping was shown to be required for balancing the Ig repertoire and to enable the orderly D to J followed by V to DJ recombination stages (Guo et al., 2011). Additional evidences supporting the relevance of CTCF in balancing the Ig repertoire came from an additional study, where they showed that CTCF depletion early during B cell development blocked differentiation, mainly by impairing both the germline transcription of the different V and J regions, and cell proliferation (Ribeiro de Almeida et al., 2011). Additionally, in this context, CTCF has been pointed as an important mediator of the IgH locus looping required for V(D)J recombination (Degner et al., 2011, Medvedovic et al., 2013, Gerasimova et al., 2015).

However the role of CTCF at later stages of B cell development has not been addressed. Given the complex transcriptional regulation involved in GC initiation, maintenance and exit and the chromatin remodeling events associated with secondary antibody diversification, we hypothesized that CTCF is very likely to play a regulatory role in these events.

5. Activation-induced deaminase

AID was first identified by Honjo's lab, as a cDNA upregulated in a B cell line (CH12F3) upon in vitro induction of CSR (Muramatsu et al., 1999). Later studies showed that AID is absolutely required for CSR and for SHM in mouse and human, where AID inactivating mutations give rise to Hyper IgM 2 immunodeficiency syndrome (Muramatsu et al., 2000, Revy et al., 2000). Given its similarity with Apobec-1, an enzyme responsible of the apolipoprotein-B mRNA editing, AID was initially hypothesized as an RNA cytidine deaminase that edits mRNA. However, several subsequent studies showed that AID deaminates cytosines present in single-strand DNA (ssDNA), but not in double-strand DNA (dsDNA), mRNA or DNA-RNA hybrids (Chaudhuri et al., 2003, Maul et al., 2011, Petersen-Mahrt et al., 2002, Pham et al., 2003, Ramiro et al., 2003, Bransteitter et al., 2003, Dickerson et al., 2003). AID initiates both SHM and CSR by deaminating dC generating deoxyuraciles (dU), thus turning C:G pairs into U:G mismatches. The final nucleotide substitution or the generation of a DSB in the case of CSR, depends on the recognition and subsequent processing of the U:G lesion. DNA replication before DNA repair causes a transition mutation from a C:G pair to T:A pair. Recognition of the U:G lesion by the mismatch repair pathway can result in the mutation of nearby A:T pairs, possibly with the contribution polymerase ζ activity. Alternatively, U can be excised by the uracyl DNA glycosilase (UNG) generating an abasic site that can be replicated by trans-lesion polymerases giving rise both transitions and transversions on the DNA (reviewed in (Di Noia and Neuberger, 2007)). On the other hand, the abasic site can also be cleaved by the apurinic/apyrimidinic (AP) endonucleases, which results in ssDNA breaks or dsDNA breaks if the abasic sites are close enough on opposite strands of the DNA, thus generating the recombination substrates for CSR. The breaks generated in the two participating switch regions are synapsed in a process that requires the participation of several proteins, including histone 2A family-member X (H2AX), p53 binding protein 1 (53BP1), ataxia telangiectasia mutated (ATM), and completed by the fusion of the two DSBs possibly by the NHEJ pathway, although this has not been unequivocally proven (reviewed in (Chaudhuri and Alt, 2004)).

5.1. Regulation of AID expression

AID activity entails the introduction of mutations and DSBs on the DNA and thus it can potentially endanger genome instability and promote oncogenic lesions (see below). A number of regulatory mechanisms are thought to contribute to an optimal balance of AID function to allow an efficient immune response while preventing collateral damage:

1. Transcriptional regulation. AID is encoded by the *Aicda* gene. *Aicda* expression is mostly, but not absolutely, restricted to activated B cells (see below). *Aicda* expression is triggered by cytokines and cell-cell interactions associated to antigen-activation of B cells during the GC reaction. Four conserved regulatory regions in the *Aicda* gene carry canonical binding sites for at least 19 transcription factors, both activators and repressors of its transcription (reviewed in (Stavnezer, 2011)). Thus, the combination of expression of the repressor proteins and transcriptional activators restricts *Aicda* expression to activated B cells.

2. Post-transcriptional regulation. The 3'-untranslated region (3'-UTR) of the AID mRNA carries binding sites for at least two miRNAs, miR-181b and miR-155. miR181b is downregulated soon after B cell activation, but increases later on during in vitro culture. miR-181b overexpression in splenic B cells downregulates AID mRNA and protein by directly binding to its 3' UTR (de Yebenes et al., 2008). On the other hand, miR-155 expression parallels AID expression in activated B cells. Mutation of the miR155-binding site in AID increased both AID mRNA and protein (Dorsett et al., 2008, Teng et al., 2008). Thus, miR181b and miR-155 seem to coordinately ensure the optimal level of AID during B cell activation.

3. Post-translational regulation. Regulation of subcellular localization, proteosomal degradation and phosphorylation all seem to contribute to the levels and activity of AID protein. AID shuttles between the nucleus and the cytoplasm (Brar et al., 2004, Ito et al., 2004, McBride et al., 2004), being actively retained in the cytoplasm and excluded from the nucleus (Methot et al., 2015). AID compartmentalization also determines protein stability, thus is destabilized in the nucleus by ubiquitin

dependent and independent pathways (Aoufouchi et al., 2008, Uchimura et al., 2011). Regarding phosphorylation, at least three positions have been defined to affect AID activity when phosphorylated, both *in vitro* and *in vivo* (McBride et al., 2008, Basu et al., 2005).

5.2. AID target specificity and lymphomagenesis

AID activity is mostly, although not exclusively, restricted to Ig genes during B cell activation. However, the mechanisms that determine the sequence specificity of AID -or targeting- remain mostly unknown. Long before the identification of AID, it was noticed that SHM preferentially accumulated on short DNA motifs in the variable region of the Ig genes that conformed to the WRCY/RGYW consensus (Dorner et al., 1998, Rogozin and Kolchanov, 1992). Later studies using biochemical assays showed that AID preferentially mutates cytosines included in WRC motifs (Bransteitter et al., 2004, Pham et al., 2003). Thus, AID activity has preference for DNA hotspots both in vitro and in vivo. However, the low complexity of AID hotspots cannot account for the preference of AID for Ig genes, suggesting that additional mechanisms are involved in defining AID target specificity.

The best-established requirement for AID activity is the transcriptional activation of the target sequence (Peters and Storb, 1996).The frequency of mutations in SHM has been shown to be proportional to the transcriptional rate in B cells (Rada and Milstein, 2001, Fukita et al., 1998, Bachl et al., 2001) and transcription of the switch region involved in CSR is associated and required for the reaction to take place(Jung et al., 1993, Stavnezer-Nordgren and Sirlin, 1986). Transcription is thought to allow the generation of ssDNA structures, either as transcription bubbles (Ramiro et al., 2003, Chaudhuri et al., 2003) or R-loops (Yu et al., 2005) that might facilitate AID accessibility to its substrate. In addition, recent studies have postulated that RNA polymerase II and the elongation factor Spt5 might be involved in the recruitment of AID to the transcribed genes (Pavri et al., 2010, Yamane et al., 2011). However, transcription is not enough to define an AID target, given the high number of transcribed genes during the GC reaction that do not accumulate AID mutations (Liu et al., 2008, Shen et al., 1998). Indeed, although a number of non-Ig genes are known to be targets

of AID (Liu et al., 2008, Pasqualucci et al., 2001), they do so at a much lower frequency -in the range of one hundredth lower- than Ig genes, which does not necessarily correlate with their transcription rate. Some recent works have proposed that AID targeting can be restricted to regions of noncoding RNA transcription that might recruit AID to these single-strand DNA-forming sites (Pefanis et al., 2014). In addition, AID recruitment seems to occur in regions with converging sense and antisense transcription (Meng et al., 2014), although these claims need to be further explored. In summary, most of the mechanisms proposed to restrict AID targeting are related directly or indirectly with transcriptional regulation.

Regardless of molecular basis of AID specificity, it is well accepted that the presumably rare off-target events of AID activity can have oncogenic impact. Specifically, AID can promote mutations and DSB followed by chromosomal translocations in other regions of the genome (Liu et al., 2008, Robbiani et al., 2009, Ramiro et al., 2006, Ramiro et al., 2004, Robbiani et al., 2008). Chromosomal translocations are the hallmark of mature B cell lymphomas, the most frequent of all human lymphomas. AID depletion delays the onset of lymphomagenesis (Kovalchuk et al., 2007, Pasqualucci et al., 2008, Ramiro et al., 2004), establishing a direct link between AID off target activity and oncogenic transformation in B cells (Figure 3).

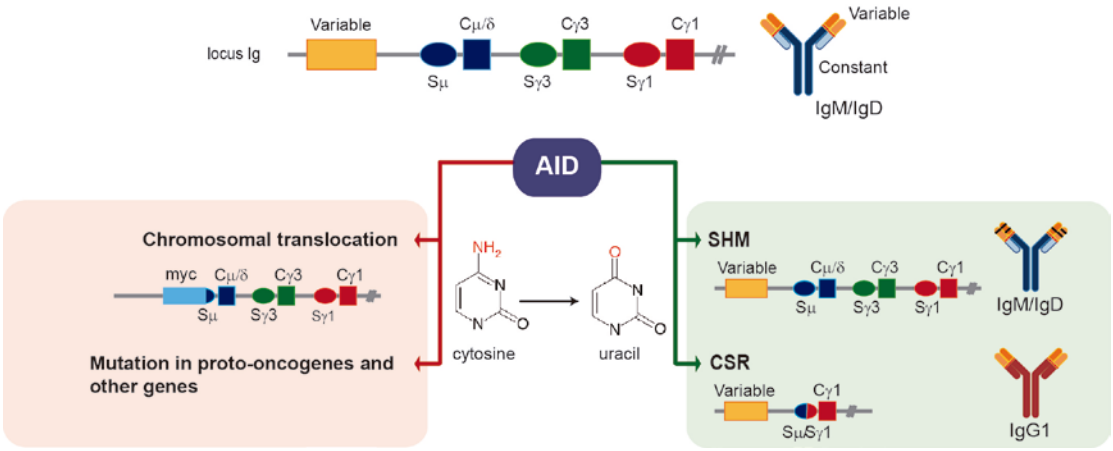

Figure 3. AID activity in Germinal Centers. In mature B cells, Ig heavy chain gene (depicted in the upper panel) is composed of a productive rearrangement of the 3' variable region (yellow square) and a 5' region that comprises a number of constant regions (small squares) each of which is preceded by a highly-repetitive non-coding switch region (small circles). The recognition of the antigen triggers a secondary diversification of the immunoglobulin in the germinal center by SHM and CSR. During SHM, point mutations (black lines) are introduced in the variable region of the Ig. CSR is a site specific recombination that generates antibodies with isotypes alternative to IgM/IgD (IgG1 is depicted in the figure). Both reactions are initiated by the deamination of citosines in the variable or the switch region of the Ig gene by AID. On the other hand, AID can also have a pathogenic activity by promoting chromosomal translocation (Myc-Ig translocation depicted), as well as mutations in proto-oncogenes and other non-Ig genes.

5.3. AID expression and activity in non-B cells

Although initially considered B cell specific, Chiba's lab provided the first evidence that AID can be induced in epithelium under inflammatory conditions (Matsumoto et al., 2007). In this study, the authors showed in humans that infection of gastric epithelium with cag+ *H. pylori* induced aberrant AID expression. In addition, in vitro TNF-α treated gastric cell lines expressed AID via NF-κB pathway. Subsequent studies carried on by this group showed AID expression in patients with chronic hepatitis or cirrhosis (Kou et al., 2007), as well as the induction of its expression in vitro by treatment with TNF-α in different cell types, such as hepatocytes, colon and cholangiocarcinoma-derived cells (Endo et al., 2007, Endo et al., 2008, Komori et al., 2008). In addition, the authors provided evidences of accumulation of nucleotide alterations in different oncogenes such as *p53* and *c-myc*, presumably triggered by AID activity.

All these works have fostered the idea that the expression and activity of AID can be broader than initially foreseen and furthermore it has been suggested that AID can link the presence of inflammation with the development of cancer (Figure 4), particularly in the epithelial context. However, this hypothesis has not been directly tested with appropriate in vivo models.

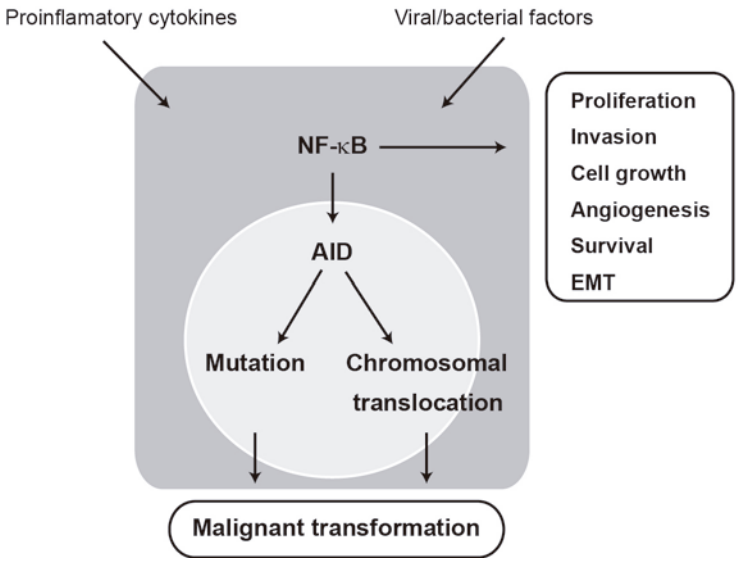

Figure 4. Connection between inflammation and aberrant AID activity in epithelium. Inflammatory cues and viral or bacterial infections might trigger AID activation through NF-κB. In a context of inflammation, where there is an increase in cell proliferation and survival, AID might promote malignant transformation through the accumulation of mutations and chromosomal translocations.

OBJECTIVES

GCs are microstructures where Ig genes diversify through SHM and CSR and their function is fundamental to generate a proper immune response. Thus, understanding the biology of this reaction, as well as the regulation of AID, has been the general objective of the present work. We defined two specific objectives:

1. To explore the involvement of CTCF in the GC reaction using a conditional knock-out mouse model.

2. To determine the role of AID in the development of carcinomas by the generation of conditional knock-in mouse models for AID overexpression in colon and pancreas.

METHODS

Mice and treatments

Mice of both genders were used unless specified otherwise. All animals were housed in the Centro Nacional de Investigaciones Cardiovasculares animal facility under a 12-h light/dark cycle with food ad libitum.

Number of animals per group to detect biologically significant effect sizes was calculated using appropriate statistical sample size formula and indicated in the biometrical planning section of the animal license submitted to the governing authority. Blinding and randomization was not applicable to the animal studies. All animal procedures conformed to EU Directive 2010/63EU and Recommendation 2007/526/EC regarding the protection of animals used for experimental and other scientific purposes, enforced in Spanish law under RD 53/2013.

Conditional mouse model for AID expression in pancreas and colon

R26AID mice were generated by insertion of a construct encoding mouse AID cDNA into the Rosa26 IRES-GFP targeting vector (Nyabi et al., 2009). AID cDNA was PCR-amplified from C57BL/6 mice (primers: forward 5'-TTCTGTGAAGACCGCAAGGCT-3'; reverse 5'-CCCTTCCCAGGCTTTGAAA-3'), cloned into the pENTR/D-TOPO Gateway vector (Invitrogen), and subsequently recombined into the Rosa 26 targeting vector, in which the cloned construct is preceded by a loxP-flanked transcriptional stop cassette and followed by an internal ribosomal entry site and GFP. The construct was linearized with PvuI before electroporation into hybrid 129/C57BL/6 ES cells. Clones positive for homologous recombination in the Rosa26 locus were identified by Southern blot of EcoRV-digested genomic DNA hybridized with a 5'-arm Rosa26 probe. R26AID mice were crossed with Villin-CRE$^{+/KI}$ (el Marjou et al., 2004) mice and p48-CRE$^{+/KI}$ (Kawaguchi et al., 2002) mice to promote expression of AID in colonic and pancreatic epithelial cells, respectively.

AID deficient mice

Balb/c $AID^{-/-}$ mice were generated by backcrossing $AID^{-/-}$ mice (Muramatsu et al, 2000) for 6 generations by speed congenics (Ramiro et al, 2004). Briefly, mice were backcrossed into the wild type Balb/c strain, genotyping each generation for 136 polymorphic markers that distinguished between Balb/c and the original C57BL/6 or CBA/J background.

DSS-induced Colitis Associated Cancer (CAC) experiments

8-10-week-old Balb/c $AID^{+/-}$ and $AID^{-/-}$ (Ramiro et al, 2004) mice were given 3% dextran sulfate sodium salt (DSS, Sigma) in their drinking water for 5 days followed by regular drinking water for 10 days. Colon samples were obtained from these mice after 10 cycles of DSS treatment, and Swiss roll preparations were fixed, embedded in paraffin for section and stained with hematoxilin/eosin.

CTCF deficient mice

Conditional knockout $CTCF^{fl/fl}$; AID-$CRE^{+/TG}$ (or controls $CTCF^{+/fl}$; AID-$CRE^{+/TG}$) mice were obtained by crossing $CTCF^{fl/fl}$ mice (Heath et al., 2008) with AID-$CRE^{+/TG}$ (Kwon et al., 2008), to allow the specific depletion of CTCF in GC B cells.

T-dependent immunization of CTCF deficient mice

T-dependent immunization was induced in 6-11 weeks $CTCF^{fl/fl}$; AID-$CRE^{+/TG}$ or $CTCF^{+/fl}$; AID-$CRE^{+/TG}$ mice by intravenously injection of 10^8 sheep red blood cells. Immunization response was analyzed in spleen seven days after injection.

Histological characterization

Immunohistochemistry

Pancreas and colon samples from aged mice were fixed in neutral-buffered 10% formalin solution (Sigma-Aldrich), embedded in paraffin blocks, and cut in 5 µM sections. Hematoxilin/eosin staining was performed using standard protocols. For immunohistochemistry, sodium citrate buffer (10mM Sodium citrate buffer (Sigma-Aldrich), pH=6) was used for antigen retrieval. Endogenous peroxidases were blocked with 3% (v/v) H_2O_2 (Merk) in Methanol (Sigma-Aldrich). The following antibodies were used: polyclonal rabbit anti-human CD3 (Dako, 1/200), rabbit anti-Ki67 (Abcam, 1/100), and biotinylated goat anti-rabbit (Abcam, 1/200). Primary antibodies were incubated overnight at 4 °C. Biotinylated secondary antibodies were incubated 1 hour at room temperature. Biotinylated antibodies were detected with the ABC system using diaminobenzidine as substrate (Vector Laboratories). Images were acquired with a Leica DM2500 microscope fitted with a 20X magnification lens. Reactive cells from 10 microscope fields per pancreas were counted using ImageJ software.

Immunofluorescence

Pancreas and colon specimens were fixed with 4% paraformaldehyde, incubated with 30% sucrose, embedded in OCT compound (Olympus) and frozen in dry ice. 10 µM sections were permeabilized (PBS 1X, 0.5% (v/v) Triton) and blocked with Image-It FX signal enhancer (Invitrogen, Molecular Probes). The following antibodies were used: rabbit anti-GFP (Abcam, 1/100), rat anti-mouse CD8α (BD Pharmigen, 1/100), goat anti-rabbit Alexa Fluor 488 (Molecular Probes, 1/500) and goat anti-rat Cy3 (Jackson ImmunoResearch, 1/500). Slides were mounted with Vectashield mount medium containing DAPI (Vector Laboratories).

Cell culture

Primary B cells

Primary B cells were obtained from $CTCF^{fl/+}$ $AID\text{-}CRE^{TG/+}$ or $CTCF^{fl/fl}$ $AID\text{-}CRE^{TG/+}$. Complete spleens were dispersed in a 70 um mesh in complete RPMI. Next, erythrocytes were lysed using erythrocyte lysis buffer (ACK Lysing Buffer, BioWhittaker) for 4 minutes at room temperature. After washing with cold complete RPMI, B cells were isolated by immunomagnetic depletion using anti-CD43 (Miltenyi Biotec). When indicated, cells were labeled with 1 µM of the cell tracer CFSE (Invitrogen). Purified B cells were cultured at a final concentration of 1.2 x 10^6 cells/ml in complete RPMI supplemented with 50 µM of 2-βMercaptoetanol (Gibco), 20 mM Hepes (Gibco). When indicated, 10 ng/ml of IL4 (PeproTech) and 25 µg/ml lipopolisacarid (LPS, Sigma-Aldrich), 25 µg/ml LPS (Sigma-Aldrich), or 10 µg/ml of anti-mouse IgM (Jackson Immunoresearch) and 1 µg/ml of anti-mouse CD40 (BD Pharmigen) were added. Cells were analyzed at different time points by flow cytometry.

Colon and pancreas cell lines

Human pancreatic and colon adenocarcinoma cell lines PaTU-8988S, AsPC-1, LoVo and SW480 cells were grown in DMEM supplemented with 10% fetal bovine serum (FBS) , HEPES (10 mM), penicillin (50 U/ml) and streptomycin (50 µg/ml). TNF-α (50 ng/ml) (Sigma-Aldrich) was added when indicated.

Pancreas explants

Primary pancreatic acinar cells were isolated and cultured as described in (Gout et al., 2013). Briefly, $R26AID^{+/KI}$; $p48\text{-}CRE^{+/KI}$ or control $R26AID^{+/+}$; $p48\text{-}CRE^{+/KI}$ mice were sacrificed by hypoxia in CO_2 camera, and pancreas was extracted and conserved in cold Hank's Balanced Saline Solution (Gibco). Complete pancreas was mechanically and enzymatically digested with collagenase IA (Sigma-Aldrich) to obtain isolated acinar structures. Acini were

grown in Waymouth's medium supplemented with 2.5% FBS, HEPES (10 mM), trypsin inhibitor (0.25 mg/ml) (Sigma-Aldrich) and recombinant human epidermal growth factor (25 ng/ml) (Sigma-Aldrich). After 24 hours of culture, cells were transferred to a collagen-coated (Type I collagen, Beckton Dickison) plate. After 48 hours, non-attached cells were removed and culture cell medium was replaced every three days.

NK killing assay in vitro

Primary pancreatic acinar cells from R26AID$^{+/+}$ p48-CRE$^{+/KI}$ and R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ mice were isolated as described above. Acinar cells were isolated from 6-days culture and individualized by trypsinization. NK cells were isolated by cell sorting from wild-type C57BL/6 mice. Both male and female aged 6 to 8 week-old were used as NK cells donors. NK cells and acinar pancreatic cells were co-cultured for 4 hours in the presence of IL2 (2000 U/ml, Peprotech) at a 1:10 (target cell: NK effector cell) ratio.

DNA damage analysis in pancreas explants

y-H2AX staining and Opera acquisition

Primary pancreatic acinar cells were grown on 96 well plates (Perkin Elmer), and y-H2AX (Millipore, 1/500) immunofluorescence was performed using standard procedures. Images were automatically acquired from each well with the Opera High-Content Screening System (Perkin Elmer). A 40x magnification lens was used and pictures were taken at non-saturating conditions. Images were segmented using DAPI staining to generate masks matching cell nuclei, and the mean per-cell y-H2AX signal was calculated.

Gene expression analysis

Messenger RNA analysis by quantitative RT-PCR

RNA was extracted from colon samples with Trizol (Invitrogen), according to manufacturer protocol. RNA from pancreas samples was extracted with GTC solution, following a standard phenol-chloroform protocol (reagents from Sigma-Aldrich), followed by DNAse treatment (Qiagen). RNA from cell pellets was extracted using RNeasy Mini Kit (Qiagen), according to manufacturer protocol, and followed by DNAse treatment (Qiagen). cDNA was synthesized from 100-600 ng of initial RNA, using Random Hexamers (Roche) and SuperScript II reverse transcriptase. mRNA was quantified by SYBR green assay (Applied Biosystems), with normalization to GAPDH. Amplification was performed in 7900HT Fast Real-Time PCR System (Applied Biosystems) equipment. Primers used are shown in table 1.

Messenger RNA analysis by Droplet digital PCR

Each 20 µl ddPCR reaction contained 10 µl of 2x ddPCR Supermix (Bio-Rad), 1 µl of cDNA generated from 1.4 µg of RNA, and 500 nM of each primer and 250 nM Taqman probe (Table 1). After droplet generation, a mean of 12,000 droplets were obtained per sample.

Samples were transferred into a 96 well plate (Eppendorf) and cycled in a thermal cycler (Bio-Rad) under the following conditions: 95 °C for 10 minutes followed by 38 cycles of 94 °C for 30 seconds, 58 °C for 1 minute and a final step at 98 °C for 10 minutes. After amplification, samples were transferred to a droplet reader (QX100 Droplet Digital PCR, Bio-Rad) from which positive-drop data were extracted with Quantasoft software.

Oligonucleotide		Sequence (5'-3')
Human-AID	Sense	AAA TGT CCG CTG GGC TAA GG
	Antisense	GGA GGA AGA GCA ATT CCA CGT
Human-GAPDH	Sense	GAA GGT GAA GGT CGG AGT C
	Antisense	GAA GAT GGT GAT GGG ATT TC
Mouse-AID	Sense	ACC TTC GCA ACA AGT CTG GCT
	Antisense	AGC CTT GCG GTC TTC ACA GAA
Mouse-GAPDH	Sense	TGA AGC AGG CAT CTG AGG G
	Antisense	CGA AGG TGG AAG AGT GGG AG
Pan-RAE	Sense	TGG ACA CTC ACA AGA CCA ATG
	Antisense	CCC AGG TGG CAC TAG GAG
Pan-RAE-FAM	probe	CCA TGA TTT ATC CGC AAA GCC AGG GCC
Elastase1	Sense	GCA CAG CAT CTT TTG TTT GGG TAA
	Antisense	GGG GAC AGT GGT CTA CTC TCT
Elastase2a	Sense	ACG AGT CCA GGA CAA TCA GAG A
	Antisense	TGA TAA GGC CAC TCA TAA AAA GGA
P53	Sense	AGC GAC TAC AGT TAG GGG GC
	Antisense	AAC CAA GTA ATC CAG AAA AAT AAG A
PRDM1	Sense	GCA AAG AGG TTA TTG GCG T
	Antisense	TGT AGA CTT CAC CGA TGA GG
CTCF	Sense	CAC CTG GGT CCT AAC AGA ACA GA
	Antisense	AGT ATG AGA GCG AAT GTG TCG TTT

Table1. Oligonucleotides used in sequencing, ddPCR and qRT-PCR.

RNA sequencing

For RNAseq library preparation, 500 ng of total RNA were used to generate barcoded RNAseq libraries using the TruSeq RNA sample preparation kit v2 (Illumina). Briefly, poly A+ RNA was purified using poly-T oligo- attached magnetic beads using two rounds of

purification followed by fragmentation and first and second cDNA strand synthesis. Next, cDNA 3' ends were adenylated and the adapters were ligated followed by PCR library amplification. Finally, the size of the libraries was checked using the Agilent 2100 Bioanalyzer DNA 1000 chip and their concentration was determined using the Qubit® fluorometer (Life Technologies). Libraries were sequenced on a HiSeq2500 (Illumina) to generate 60 bases single reads. FastQ files for each sample were obtained using CASAVA v1.8 software (Illumina).

Sequencing reads were pre-processed by means of a pipeline that used FastQC (Krueger et al., 2011), to asses read quality, and Cutadapt (Criscuolo and Brisse, 2014) to trim sequencing reads, eliminating Illumina adaptor remains, and to discard reads that were shorter than 30 bp. The resulting reads were mapped against the mouse transcriptome (GRCm38, release 76; aug2014 archive) and quantified using RSEM v1.17 (Li and Dewey, 2011). Data were then processed with a diferential expression analysis pipeline that used Bioconductor package EdgeR (Robinson et al., 2010) for normalization and diferential expression testing.

CTCF protein detection

Spleen hCD2 positive B cells were sorted from LPS/IL4 or CD3/CD28 cultures after 48h of stimulation. For CTCF immunoblotting, cell pellets were incubated on ice for 20 min in RIPA lysis buffer in the presence of protease inhibitors (Roche) and lysates were cleared by centrifugation. Total protein was sized-fractioned by SDS-PAGE on 8% acrylamide-bisacrylamide gel and transferred to Protan nitrocellulose membrane (Whatman) in transfer buffer (0.19M glycine, 25 mM Tris base and 0.01% SDS) containing 20% methanol. Gels were transferred for 45 min at 0.4A. Membranes were probed with anti-mouse CTCF (1/2500, Bethyl laboratories) and anti-mouse-tubulin (1/5000, Sigma-Aldrich). Then, membranes were incubated with corresponding HRP-conjugated antibodies (Dako) and developed with Amesham ECL Western Blotting Detection Reagent (GE Healthcare Life Sciences).

DNA Sequencing

Sanger sequencing

DNA was extracted from pancreatic tissue following an standard phenol:chloroform:Isoamyl Alcohol (25:24:1) protocol. Precipitation was performed in ethanol, using the Pellet Paint co-precipitant (Novagen), according to manufacturer protocol.

Amplification was carried out with 2.5 U of Pfu Ultra (Stratagene) in a 50 µl reaction, containing 200 ng of genomic DNA, following the profile: 94 °C- 5 minutes, 25 cycles a 94 °C-10 seconds, 60 °C-30 seconds, y 72 °C-1 minute. Primers used are shown in Table 1. The amplified fragment was isolated from 1% agarose gel and purified in column (*illustra GFXTM Gel Band Purification Kit*), following manufacturer protocol. Purified fragment was cloned in the *pGEM-T Easy vector* and sequenced using the SP6 primer. The sequences obtained were analyzed for the presence of mutations with the program *Seqman (Lasergen)*.

Next generation sequencing for detection of mutations

DNA from three R26AID$^{+/+}$; p48-CRE$^{+/KI}$ and three R26AID$^{+/KI}$; p48-CRE$^{+/KI}$ was extracted and amplified following the conditions used in Sanger sequencing. Primers used for amplification are depicted in Table 1. After amplification, equimolar quantities of DNA from each mouse were pooled. The amplification products were fragmented in 100-300 bp fragments using Covaris S2 shearing instrument. Specific adapters were added to each sample to its identification. Next, samples were purified using the Agencourt AMPure XP (Beckman Coulter Genomics) kit, following manufacturer instructions. To generate the libraries, 10 additional cycles of amplification were performed using specific oligonucleotides provided by Illumina, with the Phusion High-Fidelity polymerase (Finnzymes). Purified libraries were sequenced in Genome Analyzer IIx (Illumina), using the SBS TruSeq v5 reactives, following manufacturer instructions. Lectures with a quality Q<20 were excluded.

The good-quality sequences were aligned Novoalign (Novocraft technologies) and were processed with SAMtools (Li et al., 2009). In order to obtain a table in excel format with all the mutations obtained, the results were parsed with a specific program designed in Perl.

Flow cytometry

Lymphoid tissues analysis in CTCF deficient mice

$CTCF^{fl/fl}$; AID-$CRE^{+/TG}$ and $CTCF^{+/fl}$; AID-$CRE^{+/TG}$ were sacrificed by hypoxia in CO_2 camera and Peyer's Patches, spleen and femur were extracted, and conserved in RPMI-1640 (Sigma-Aldrich) supplemented with 10% FBS and penicillin (50 U/ml) and streptomycin (50 µg/ml) (thereafter complete RPMI). Peyer's patches and spleens were dispersed in a 70 mm mesh in complete RPMI and cell suspension obtained were centrifuged (10'/400 x g/ 4°C). Bone marrow cells were obtained by perfusion with complete RPMI through femur channel. The spleen and bone marrow cell suspensions obtained were centrifuged and erythrocytes were lysed using erythrocyte lysis buffer (ACK Lysing Buffer, BioWhittaker) for 4 minutes at room temperature. After lysis, cells were washed with complete RPMI and resuspend in PBS 1X-2% FBS.

Cells obtained from Peyer's Patches, spleens and bone marrows were stained with antibodies conjugated to different fluorochromes or to biotin. Cells were centrifuged (10 min/400xg/4°C) and resuspended in PBS 1X- 2% FBS and the antibody combination indicated in each case. Cells were incubated on ice for 20 min, washed with PBS 1X- 2% FBS and centrifuged (10 min/400xg/4°C). When direct antibodies were used, cells were resuspend in PBS 1X- 2% FBS for its analysis by flow cytometry. When biotin-conjugated antibodies were used, cells were incubated in a second step with streptavidin conjugated to the corresponding fluorchrome in PBS 1X- 2% FBS, as indicated before. Next, cells were washed with PBS 1X- 2% FBS, centrifuged (10 min/400xg/4°C) and resuspend in PBS 1X- 2% FBS. Cells were stained with the appropriate combination of the antibodies shown in table 2. Data were acquired in LSRFortessa (BD Biosciences) and analyzed with the FlowJo software.

Cell cycle analysis in CTCF deficient primary B cells

Primary B and T cells cultured for different time points were harvested, washed with PBS 1X-2% FCS, and incubated with Hoescht during 1 hour at 37°C. After incubating with Hoescht, membrane was stained with the appropriated combination of antibodies (Table 2). Data were acquired in LSRFortessa (BD Biosciences) and analyzed with the FACSDiva software (BD Biosciences).

Pancreas explant flow cytometry

RAE detection: Acinar cells were isolated and individualized from 6-days explant cultures. Cells obtained from the cultures were centrifuged (3 min/300xg/4°C) and resuspended in HBSS 1X- 2% FBS and the anti-mouse RAE purified antibody. Cells were incubated on ice for 20 min, washed with HBSS 1X- 2% FBS and centrifuged (3 min/300xg/4°C). Cells were incubated in a second step with secondary antibody conjugated to Alexa Fluor-647 (Life Technologies) in HBSS 1X- 2% FBS, as indicated before. Next, cells were washed with HBSS 1X- 2% FBS, centrifuged (3 min/300xg/4°C) and resuspend in HBSS 1X- 2% FBS.

Analysis of NK killing assay: After 6 days of culture, acinar cells were trypsinized and stained with CFSE, in order to distinguish them from NK cells. Acinar cells were cultured in presence of NK cells for 4 hours. Reaction was stopped by introducing culture plates on ice. Total cells were directly stained with DAPI, to detect the presence of death cells. Data are presented as the proportion of $CFSE^{+}DAPI^{+}$ cells normalized to the same population in cultures lacking NK cells.

Statistics

Statistical analyses were performed with GraphPad Prism (version 6.01 for Windows, GraphPad Software, San Diego, CA, USA) using two-tailed Student's t-test. $P \leq 0.05$ was

considered statistically significant. Error bars in figures represent standard error of the mean (SEM). Normal distribution of data was assessed by applying a D'Agostino & Pearson omnibus normality test. F-test was used to compare variances between groups. For the survival analyses, GraphPad Prism was used and the Mantel–Cox test was applied. Differences were considered statistically significant at $P \leq 0.05$. For the correlation analysis performed, D'Agostino & Pearson omnibus normality test was applied. As samples did not follow a Gaussian distribution, non-parametric Spearman correlation analysis was performed.

Fluorochrome-conjugated antibodies

Antibody	Company	Fluorochrome
B220 (CD45R)	BD Horizon	Brilliant Violet
CD21	BD Biosciences	FITC
CD93	eBioscience	APC
CD5	BD Biosciences	FITC
IgD	BD Biosciences	FITC
IgM	Invitrogen	PE
CD25	BD Biosciences	APC
Fas	BD Biosciences	PE
GL7	BD Biosciences	FITC

Biotin-conjugated antibodies

Antibody	Company
CD23	BD Biosciences
CD43	BD Biosciences
CD19	BD Biosciences
IgA	BD Biosciences
IgG1	BD Biosciences

Streptavidin

Fluorochrome	Company
PE-Cy7	BD Biosciences

Table 2. Antibodies used in FACS analysis

RESULTS

1. Analysis of the role of CTCF during the Germinal Center reaction

Transition through the different stages of the GC and PC differentiation involves the coordinated expression of gene networks. Bcl-6 is the master regulator of the GC reaction and acts as a transcriptional repressor to control genes involved in activation, survival, DNA-damage response and cell cycle arrest, among others pathways (Reviewed in (Basso and Dalla-Favera, 2010)). Bcl-6 also plays an important role in preventing plasma cell differentiation by the repression of Blimp-1. In turn, Blimp-1 was firstly described as a factor sufficient to drive plasma cell differentiation (Turner et al., 1994) and is considered the key factor controlling transition from GC B cells to antibody-secreting plasma cells (Shapiro-Shelef et al., 2003, Shaffer et al., 2002), by its ability to repress cell proliferation and to establish the PC program. CTCF is a ubiquitously expressed chromatin remodeler and proposed to have a general role in the control of gene transcription (reviewed in (Ong and Corces, 2014)). In B lymphocytes, a number of studies have shown the relevance of CTCF during B cell development. In particular, CTCF is a key regulator of VDJ recombination and B cell differentiation, by promoting the interaction between different V exons (Garrett et al., 2005, Degner et al., 2011, Guo et al., 2011, Ju et al., 2011, Volpi et al., 2012, Medvedovic et al., 2013, Lin et al., 2015, Ribeiro de Almeida et al., 2011, Ribeiro de Almeida et al., 2012). However, the role of CTCF in mature B cells, and particularly in the context of B cell activation in GCs, has not been addressed.

1.1. $CTCF^{fl}$-AIDCRE TG B cells develop normally

To address the role of CTCF during the germinal center (GC) reaction we specifically depleted CTCF in GC B cells by crossing mice carrying a floxed CTCF allele ($CTCF^{fl}$ mice) with transgenic AID-CRE mice ($AID\text{-}CRE^{TG/+}$), where the Cre recombinase is inserted together with a cDNA encoding the human CD2 molecule (hCD2) in a BAC construct that contains the complete AID locus (Figure 5A). Thus, in $CTCF^{fl/fl}$;$AID\text{-}CRE^{TG/+}$ mice ($CTCF^{fl/fl}$ hereafter for

simplicity) the expression of CRE and the deletion of CTCF is induced concomitantly with the expression of AID -therefore specifically expressed in GC B cells- and can be followed by the surface expression of the hCD2 (Figure 5B). In all the experiments $CTCF^{fl/+}$ $AID\text{-}CRE^{TG/+}$ mice were used as control ($CTCF^{fl/+}$ mice).

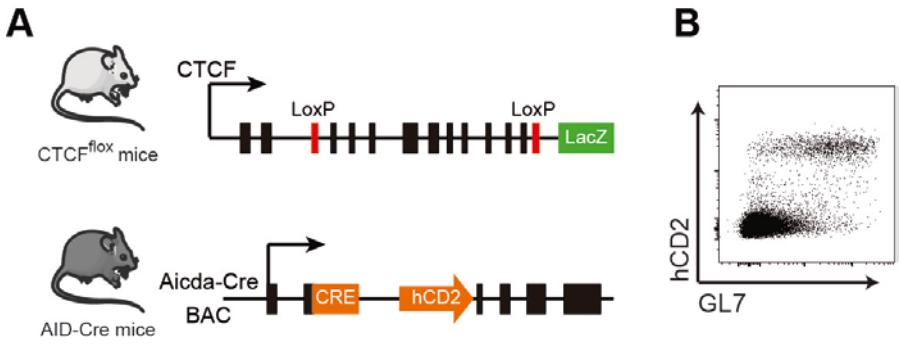

Figure 5. Conditional depletion of CTCF in GC B cells. Mouse model used to specifically remove CTCF in GC B cells. A) Schematic of the constructs used for conditional depletion of CTCF in GC B cells. Mice carrying the CTCF allele flanked by LoxP sites (top) were crossed with transgenic AID-CRE mice (bottom), where a cassette coding for Cre recombinase and the hCD2 receptor is inserted in an AID-BAC. B) Analysis of hCD2 expression in GCs B cells ($GL7^{+}$) in Peyer's Patches from $CTCF^{fl/+}$ mice.

We first analyzed B cell differentiation in bone marrow and the subsequent maturation that takes place in spleen, where immature B cells differentiate into transitional (T), marginal zone (MZ) and follicular (FO) mature B cells, of $CTCF^{fl/fl}$ and $CTCF^{fl/+}$ control mice (Figure 6, 7). The analysis of B cell populations both in bone marrow and spleen did not showed any difference between CTCF deficient and control mice. We can conclude that the depletion of CTCF promoted by AID-Cre recombinase does not affect B cell differentiation.

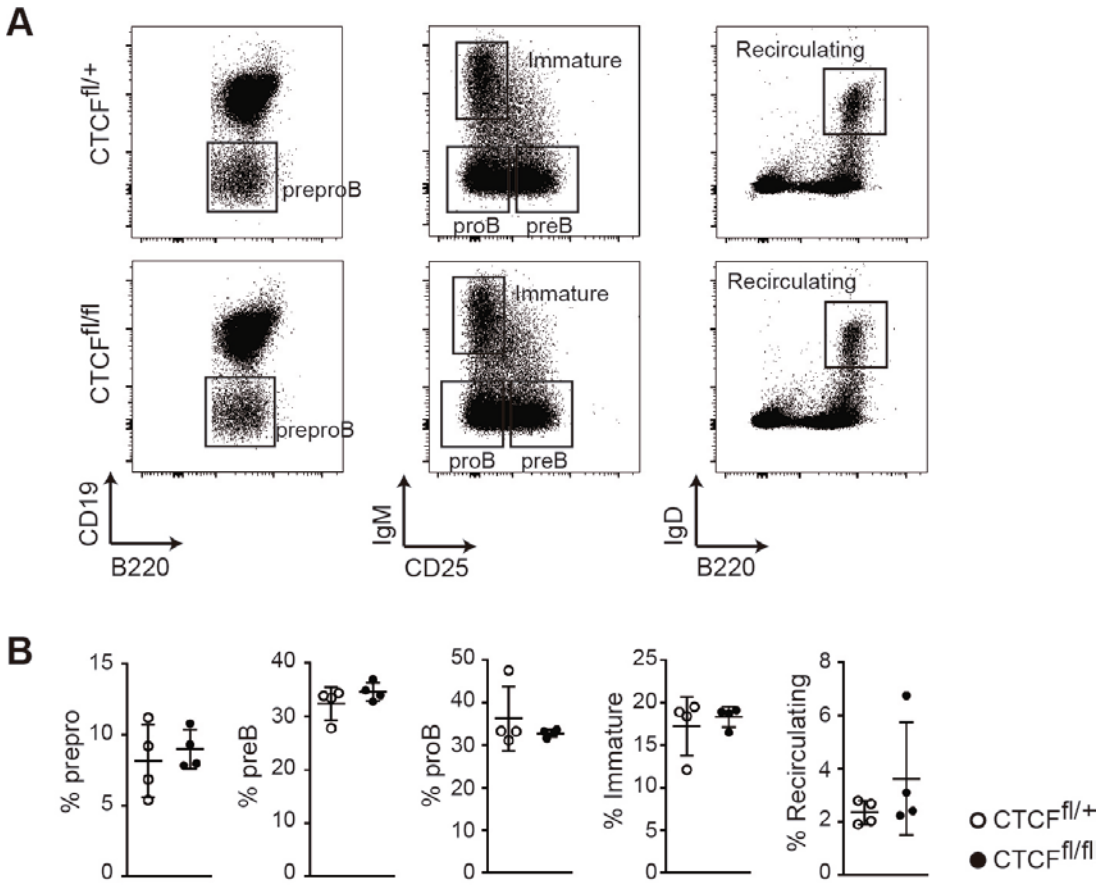

Figure 6. CTCF depletion in GC B cells does not affect B cell differentiation in the bone marrow. B cell differentiation was analyzed by FACS in the bone marrow from $CTCF^{fl/fl}$ and $CTCF^{fl/+}$ control mice. A) Quantitative FACS analysis of B cell differentiation in bone marrow. The different populations were defined as: preproB (B220+ CD19-); preB ($B220^{+}CD19^{+}CD25^{+}IgM^{-}$); proB ($B220^{+}CD19^{+}CD25^{-}IgM^{-}$); Immature ($B220^{+}CD19^{+}CD25^{-}IgM^{+}$); Recirculating ($B220^{+}CD19^{+}IgD^{+}$). B) Quantification of all the populations shown in (A), measured as the percentage of each population referred to total $B220^{+}$ cells. All data are values +/- Standard Deviation ($CTCF^{fl/+}$, white dots; $CTCF^{fl/fl}$; black dots). Statistical analysis was done with two-tailed unpaired Student's t test.

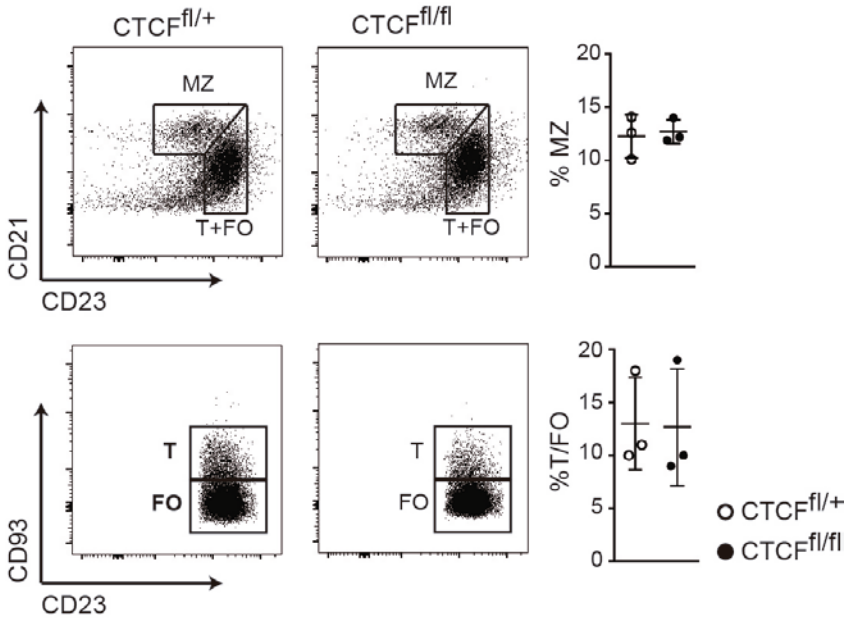

Figure 7. CTCF depletion in GC B cells does not affect B cell differentiation in the spleen. B cell differentiation was analyzed by FACS in the spleen from $CTCF^{fl/fl}$ and $CTCF^{fl/+}$ control mice. The different populations were defined as: Marginal Zone (MZ: $B220^{+}CD21^{++}CD23^{+}$); Follicular (FO: $B220^{+}CD21^{+}CD23^{++}CD93^{-}$); Transitional (T: $B220^{+}CD21^{+}CD23^{++}CD93^{-}$). Quantifications are shown on the right, measured as the percentage of each population referred to total $B220^{+}$ cells. All data are values +/- Standard Deviation ($CTCF^{fl/+}$, white dots; $CTCF^{fl/fl}$; black dots). Statistical analysis was done with two-tailed unpaired Student's t test.

1.2. CTCF is required for the GC reaction

To evaluate the impact of CTCF depletion in GC B cells we first analyzed the constitutive GC response in the Peyer's Patches (PPs) of $CTCF^{fl/fl}$ and $CTCF^{fl/+}$ control mice (Figure 8). We measured the presence of Fas+GL7+ B cells by FACS, and found that in CTCF deficient mice the GC B cell population is dramatically reduced (3.8 ± 0.6% in $CTCF^{fl/+}$ mice versus 0.57±0.1% in $CTCF^{fl/fl}$ mice) (Figure 8, upper plots). In addition, we observed that CTCF deficient mice lack almost completely IgA switched B cells (3.88±0.4% in $CTCF^{fl/+}$ mice versus 1.03±0.2% in $CTCF^{fl/fl}$ mice) (Figure 8, middle plots). The analysis of the Cre reporter gene hCD2 showed a clear population of hCD2 positive cells in control mice, whereas this

population was almost absent in CTCF deficient mice (5.48±0.5% in $CTCF^{fl/+}$ mice versus 0.8±0.2% in $CTCF^{fl/fl}$ mice) (Figure 8, bottom plots). These results suggest that CTCF is required for the GC reaction in vivo.

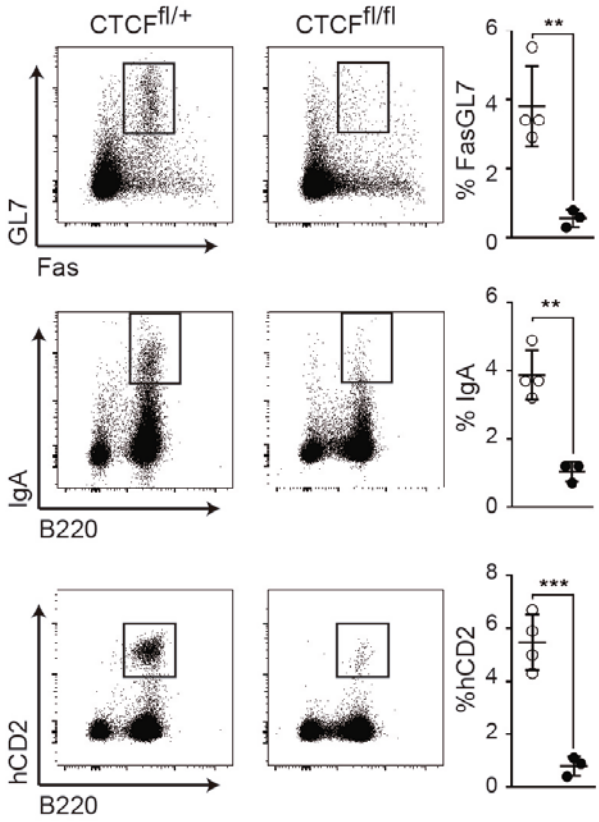

Figure 8. CTCF is required for the constitutive GC reaction. Quantitative FACS analysis of GL7, Fas, IgA and hCD2 expression in Peyer's Patches B cells from $CTCF^{fl/+}$ (n=4) and $CTCF^{fl/fl}$ (n=3) mice. Top: $Fas^{+}GL7^{+}$ B cells. Middle: $B220^{+}IgA^{+}$ total live lymphocytes. Bottom: $B220^{+}hCD2^{+}$ total live lymphocytes. Quantifications are shown on the right, measured as percentage of each population referred to total $B220^{+}$ cells. p(FasGL7)= 0.0056; p(IgA)= 0.0015; p(hCD2)= 0.0008. All data are values +/- Standard Deviation ($CTCF^{fl/+}$, white dots; $CTCF^{fl/fl}$; black dots). Statistical analysis was done with two-tailed unpaired Student's t test.

To determine the role of CTCF in *de novo* GC, we immunized $CTCF^{fl/fl}$ and $CTCF^{fl/+}$ mice with sheep red blood cells (SRBC) and analyzed the GC response in spleen 7 days after

immunization (Figure 9A). Quantification of GC B cells (Fas^+GL7^+) and IgG1-switched cells by FACS showed that both populations were reduced in the absence of CTCF (FasGL7: 3.03 ± 0.4 in $CTCF^{fl/+}$ mice versus 0.1 ± 0 $CTCF^{fl/fl}$ mice; IgG1: 0.42 ± 0.06 in $CTCF^{fl/+}$ mice versus 0.02 ± 0.02 in $CTCF^{fl/fl}$ mice). In addition, immunohistochemistry staining revealed that CTCF deficient mice lack Bcl6 positive B cells (Figure 9B), a marker of GC B cells. Together, these results showed that CTCF is required for both constitutive and *de novo* GC reactions.

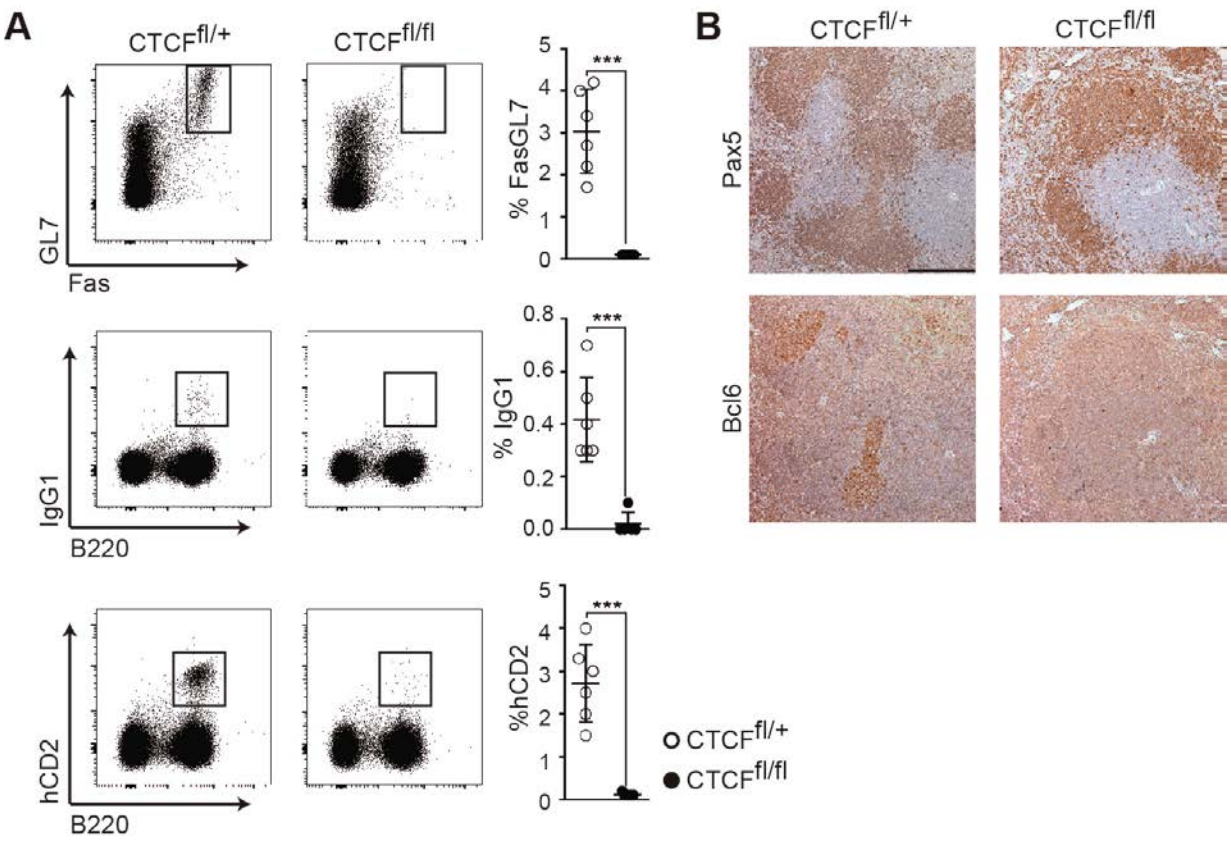

Figure 9. CTCF is required for de novo GC reaction. A) FACS analysis of GL7, Fas, IgG1 and hCD2 expression in splenic B cells from $CTCF^{fl/+}$ (n=6) and $CTCF^{fl/fl}$ (n=5) mice 7 days after SRBC immunization. Top: Fas^+GL7^+ B cells. Middle: $B220^+IgG1^+$ total live lymphocytes. Bottom: $B220^+hCD2^+$ total live lymphocytes. Quantifications are shown on the right, measured as percentage of each population referred to total $B220^+$ cells. p(FasGL7)= 0.0001; p(IgG1)= 0.0005; p(hCD2)= 0.0001. B) Representative Pax5 (top) and Bcl6 (bottom) immunohistochemistry of spleen from $CTCF^{fl/+}$ and $CTCF^{fl/fl}$ mice after 7 days of SRBC immunization. All data are values +/- Standard Deviation ($CTCF^{fl/+}$,

white dots; $CTCF^{fl/fl}$; black dots). Statistical analysis was done with two-tailed unpaired Student's t test.

1.3. CTCF depletion differentially affects B cells under different stimuli conditions

To address the molecular mechanisms of GC impairment in CTCF deficient mice we made use of an standard in vitro system in which LPS/IL4 stimulation of naïve B cells allows recapitulating many features of the GC reaction such as induction of B cell proliferation, AID expression, γ 1 switch region transcription and CSR . Splenic B cells coming from $CTCF^{fl/fl}$ or control mice were stimulated with LPS/IL4 and the B cell activation was tracked by the expression of the hCD2 reporter. Surprisingly, we did not find any difference in the proportion of hCD2 positive cells between CTCF deficient and control mice (82.84±2.4% in $CTCF^{fl/+}$ cells and 84.31±3.8% in $CTCF^{fl/fl}$ cells) (Figure 10A, B). To rule out that the absence of phenotype could be due to a defective deletion of the CTCF allele, we first measured CTCF protein by WB in sorted $hCD2^{+}$ B cells from LPS/IL4 cultures (Figure 10C). We observed a reduction of more than 80% of CTCF protein in $CTCF^{fl/fl}$ mice. In addition, we measured CTCF mRNA levels, both by qRT-PCR and RNA-sequencing, and observed a significant reduction in $CTCF^{fl/fl}$ compared with control mice (Figure 10D). The discrepancy observed between in vivo and in vitro effect after CTCF depletion suggests that LPS and IL4 cannot recapitulate faithfully GC B cell activation *in vivo* and bypasses the need of activated B cells for CTCF.

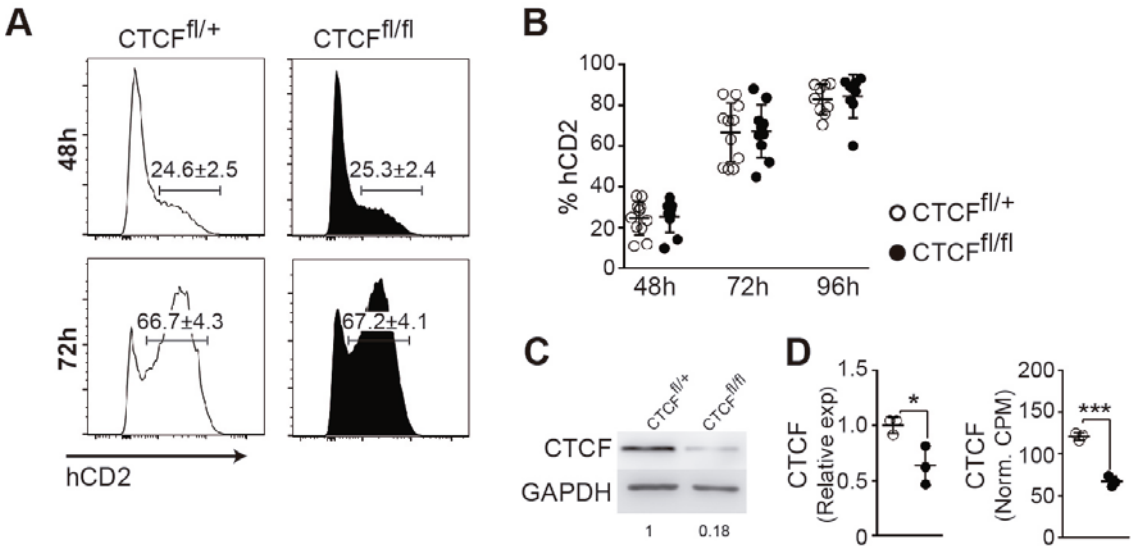

Figure 10. CTCF is not required for LPS/IL4-mediated B cell activation in vitro. In vitro activation of splenic B cells from control and CTCF deficient mice was analyzed in LPS/IL4 for different time points. A) Analysis of $hCD2^{+}$ in LPS/IL4 stimulated splenic B cells from control and CTCF deficient mice at 48h and 72h. B) Percentage of $hCD2^{+}$ cells at the different time points analyzed. n($CTCF^{fl/+}$)=11; n($CTCF^{fl/fl}$)=10. C) CTCF Western Blot in $hCD2^{+}$ purified B cells from control and CTCF deficient mice after 72 hours of stimulation with LPS/IL4. CTCF amount normalized to GAPDH is shown. D) CTCF mRNA quantification by qRT-PCR (p=0.0283) and by RNA-seq (p=0.0003). All data are values +/- Standard Deviation ($CTCF^{fl/+}$, white dots; $CTCF^{fl/fl}$; black dots). Statistical analysis was done with two-tailed unpaired Student's *t* test.

To address this issue, we sought for *in vitro* stimulation conditions that could reproduce more closely B cell activation in vivo. B cells can be stimulated in vitro by T cells in the presence of immobilized anti-CD3 together with soluble anti-CD28 (Klaus et al., 1994, Johnson-Leger et al., 1998, Klaus et al., 1999). To allow B cell stimulation in the absence of LPS/IL4, we co-cultured them with CD3/CD28 stimulated T cells in a 1:1 ratio. When we measured the levels of hCD2 to track B cell activation we observed that after 72 and 96 hours of stimulation, the population of hCD2 positive cells was significantly reduced in the CTCF deficient samples compared with control mice (72h: 17.34 ± 2.7 in $CTCF^{fl/+}$ cells versus 10.71 ± 1.09 in $CTCF^{fl/fl}$ cells; 96h: 23.67 ± 2.7 in $CTCF^{fl/+}$ cells versus 15.33 ± 1.4 in $CTCF^{fl/fl}$ cells) (Figure 11A, B). Quantification of CTCF protein in $hCD2^{+}$ cells by WB revealed a similar

level of depletion to the observed in LPS/IL4 cultures (Figure 11C), which agreed with a significant reduction in the CTCF mRNA level, both by qRT-PCR and RNA-sequencing (Figure 11D). Thus, despite similar levels of CTCF depletion, B cell stimulation by LPS and IL4 is refractory to CTCF loss, while B cell activation in the presence of T cells requires CTCF. These results indicate that the stimulation pathway might determine the susceptibility of B cells to CTCF deficiency, and that T-B co-culture might better recapitulate B cell stimulation in GCs.

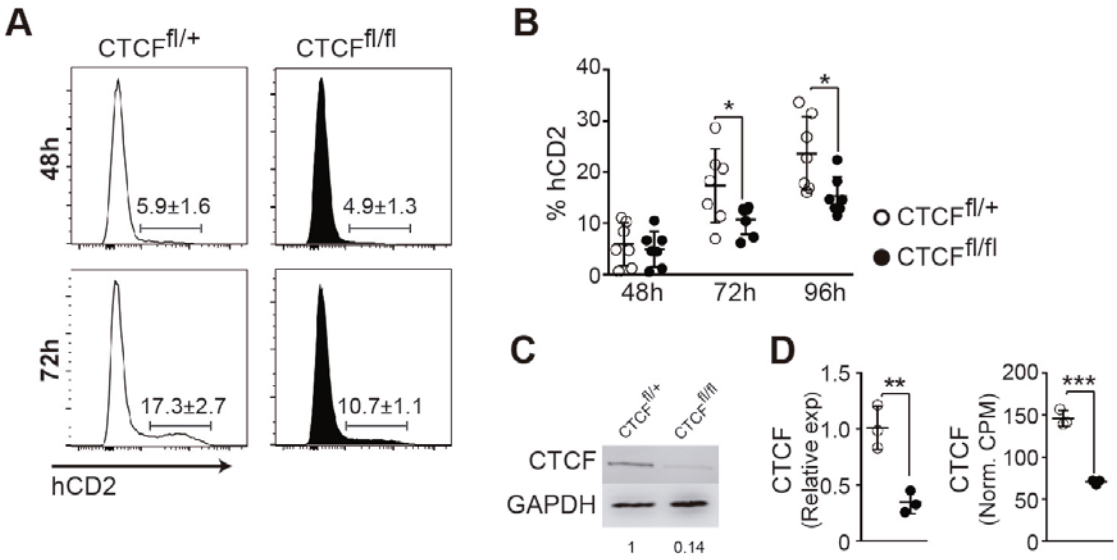

Figure 11. **CTCF is required for CD3/CD28 activated B cells in T-B co-culture.** A) Analysis of $hCD2^+$ in CD3/CD28 T-B co-cultures from control and CTCF deficient mice at 48h and 72h. B) Percentage of $hCD2^+$ cells at the different time points analyzed. n($CTCF^{fl/+}$)= 7;n($CTCF^{fl/fl}$)= 7. p(72h)= 0.0438; p(96h)= 0.0187. C) CTCF Western Blot in $hCD2^+$ purified B cells from control and CTCF deficient mice after 72 hours of stimulation with CD3/CD28. Normalized CTCF amount relative to GAPDH is shown. D) mRNA CTCF quantification by qRT-PCR (p= 0.0059) and by RNA-seq (p= 0.0002). All data are values +/- Standard Deviation ($CTCF^{fl/+}$, white dots; $CTCF^{fl/fl}$; black dots). Statistical analysis was done with two-tailed unpaired Student's *t* test.

1.4. In vitro T-B co-culture better mimics in vivo stimulation

To understand the differential requirement for CTCF of B cells under different activation cues we analyzed the transcriptional profile of CTCF proficient B cells stimulated with LPS and IL4 and B cells stimulated with CD3/CD28 T cells and compared them with in vivo activated GC B cells. For that, we performed RNA-seq of isolated $Fas^{+}GL7^{+}$ GC B cells from PPs, of $hCD2^{+}$ B cells isolated from CD3/CD28 T cell cultures or $hCD2^{+}$ cells isolated from LPS/IL4 cultures. Each of these samples was compared with RNA-seq from resting splenic B cells in triplicates and the three pairwise comparisons were first analyzed with Venn diagrams (Figure 12). We observed that all the three stimulation conditions shared a high proportion of both downregulated and upregulated genes, suggesting that in vitro activation conditions properly recapitulate many of the events that take place in vivo. In addition, we found that CD3/CD28 T cell stimulated and LPS/IL4 stimulated B cells shared 12% of their transcriptional programs. However, we observed that CD3/CD28 T stimulated B cells and GC B cells shared a specific set of 267 downregulated genes (8.89%, shaded area 1), whereas LPS/IL4 stimulated B cells only shared 97 unique downregulated genes (3.29% shaded area 2) with GC cells (Figure 12A). Similar results were found for upregulated genes (Figure 12B).

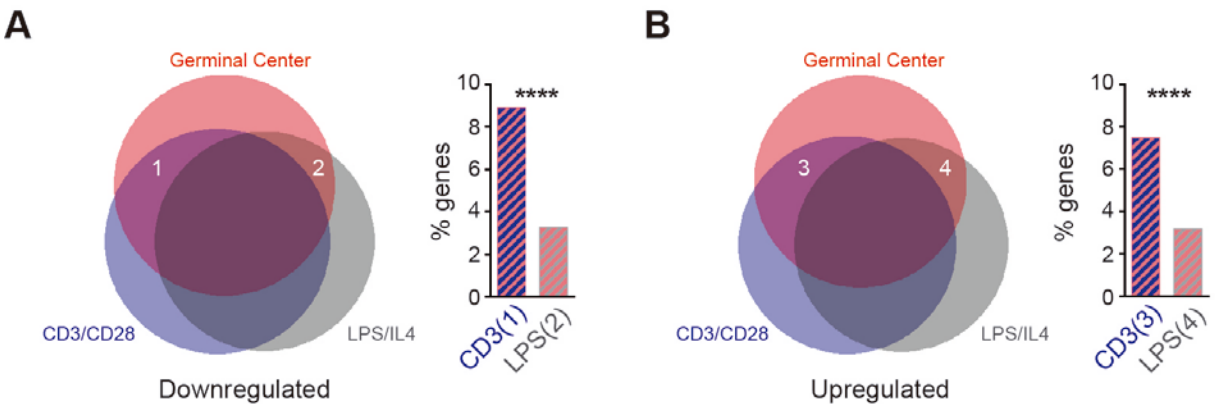

Figure 12. In vitro B cell stimulation through T cells recapitulates the GC reaction. RNA-seq analysis was performed in GC B cells from Peyer's Patches (GC) and $hCD2^+$ sorted B cells from WT mice stimulated for 48h in the CD3/CD28 T-B co-culture (CD3) or LPS/IL4 (LPS) and data were compared pairwise with those obtained from resting B cells. Overlaps of downregulated (A) or upregulated (B) genes across all three conditions are depicted in Venn diagrams. Percentage of the common downregulated (A) or upregulated (B) genes between GC and CD3/CD28 or between GC and LPS/IL4 (shaded areas 1-4) is shown. Statistical analyses were done with two-tailed chi-square analysis. p-value < 0.00001.

In order to analyze in more detail the differences between stimulations, we selected the 10% most differentially expressed genes between GC and resting B cells, and plotted the Z-score of these genes in all stimulation conditions (resting, LPS/IL4 stimulated, GC and CD3/CD28 T stimulated B cells). Importantly, this analysis revealed that the highest similarity between GC and CD3/CD28 stimulated cells occurred in genes related with cell cycle, a key feature of the GC reaction (Figure 13A). This concept is reflected in the analysis of individual genes (Figure 13B).

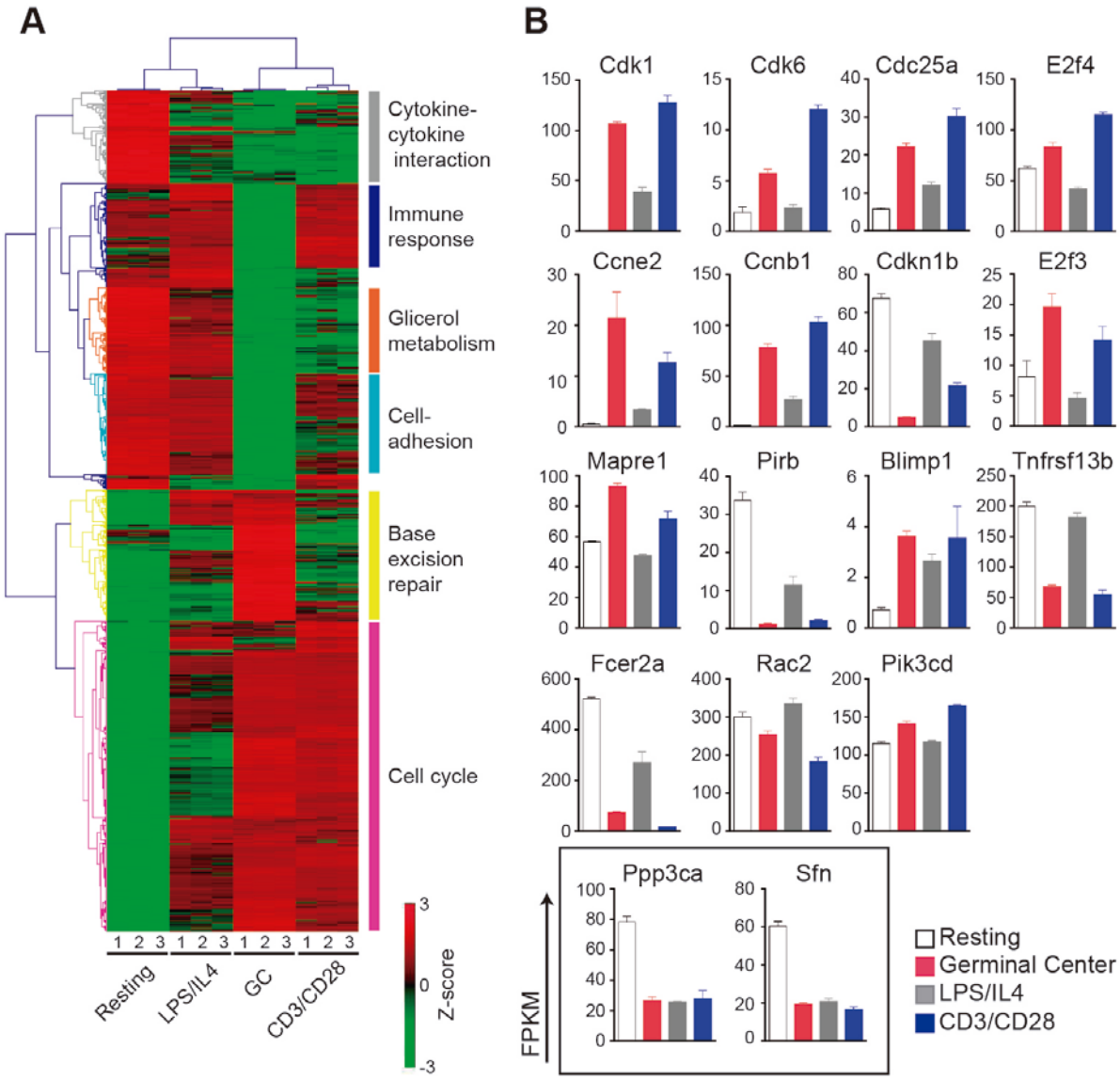

Figure 13. Transcriptional comparison of GC B cells, LPS/IL4 and CD3/CD28 stimulated B cells. A) Heatmap of 10% more differentially expressed genes between resting and GC cells. The scale bar indicates z-scores of gene expression values for resting, GC and LPS/IL4 and CD3/D28 stimulated B cells. Clustering was performed using the average linkage method based on Pearson correlation distance. Each column represents an independent replicate. Pathway enrichment analysis of gene clusters is shown on the right. B) Expression profile of selected genes. Unchanged genes are represented in the square.

These data suggest that co-culturing B cells together with T cells recapitulates the in vivo GC transcriptional program better than the standard LPS/IL4 stimulation, which in turn could explain the differential requirement for CTCF of B cells stimulated under these conditions.

1.5. CTCF transcriptionally regulates key processes of GC biology

To further characterize the role of CTCF in the GC reaction, and given the unfeasibility to obtain CTCF deficient cells in vivo, we decided to analyze the transcriptional differences between CTCF deficient and control mice in both in vitro stimulation conditions. To that end, we performed RNA-seq analysis of hCD2 positive splenic B cells from CD3/CD28 T or LPS/IL4 cultures. Comparison of CTCF deficient and proficient cells from CD3/CD28 T cultures, revealed differentially expressed 2229 genes (adjusted p-value<0.05) of which approximately half were upregulated (1154) and half were downregulated (1064) (Figure 14A). In sharp contrast, only 50 genes were differentially expressed in CTCF deficient versus proficient cells from LPS/IL4 cultures (Figure 14A). This notion was reinforced when analyzing the Z-scores of the 20% most differentially expressed genes between CTCF deficient and control CD3/CD28 T stimulated cells (Figure 14B, Annex). Together, these data agree with the finding that CTCF is dispensable for LPS/IL4 stimulated B cells but required for CD3/CD28 T stimulated B cells and for GC B cells.

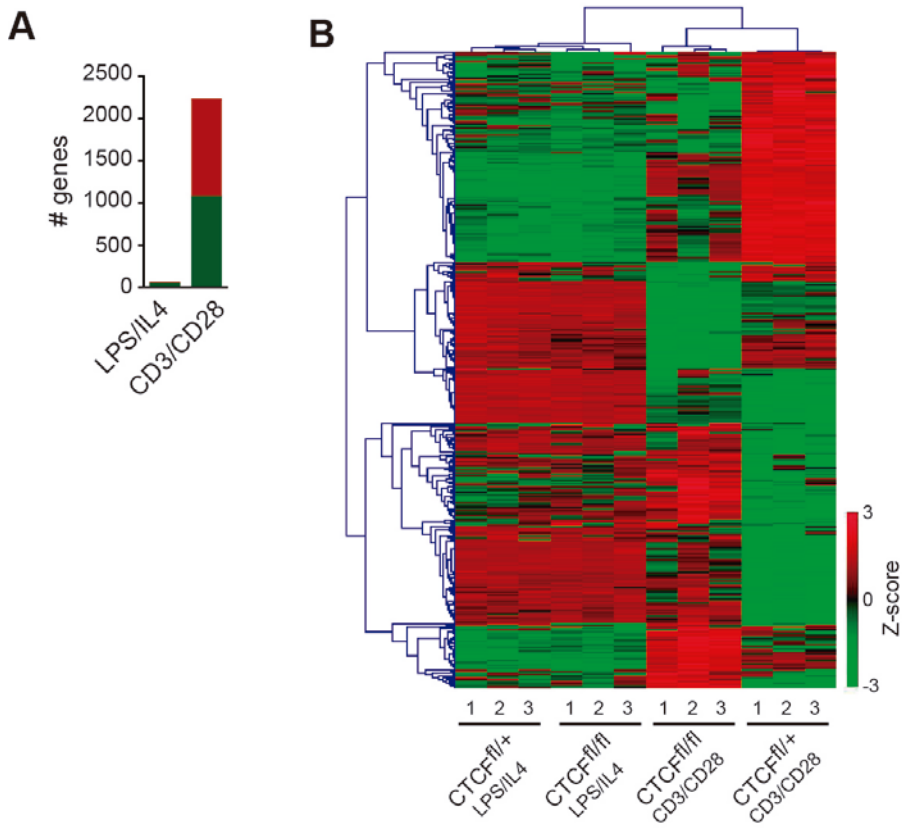

Figure 14. CTCF deficiency differentially affects B cells under different stimulation conditions. RNA-seq analysis was performed in hCD2^{+} sorted B cells stimulated for 48h in the presence of CD3/CD28 T or LPS/IL4. A) Number of genes differentially transcribed between CTCF$^{fl/fl}$ and control samples in LPS/IL4 or CD3/CD28 conditions. Red: upregulated genes; green: downregulated genes. B) Heatmap of the 20% genes most differentially expressed between CTCF$^{fl/fl}$ and CTCF$^{fl/+}$ B cells after CD3/CD28 T stimulation. The scale bar indicates z-scores of gene expression values for CTCF$^{fl/fl}$ or CTCF$^{fl/+}$ samples in LPS/IL4 and CD3/CD28 stimulation conditions. Clustering was performed using the average linkage method based on Pearson correlation distance. Each column represents an independent replicate.

The functional relevance of the genes affected by CTCF depletion is revealed by the finding that most of the genes controlled by CTCF (75,7%, CTCF$^{fl/fl}$ versus CTCF$^{fl/+}$ CD3/CD28 T stimulated B cells) are part of the GC transcriptional program in vivo (Figure 15A). Importantly, we observed that almost 65% of these common genes (1095 out of 1688) negatively correlated when comparing fold change expression (Figure 15B). This result suggests that CTCF might be involved in the regulation of a significant number of genes involved in the GC reaction.

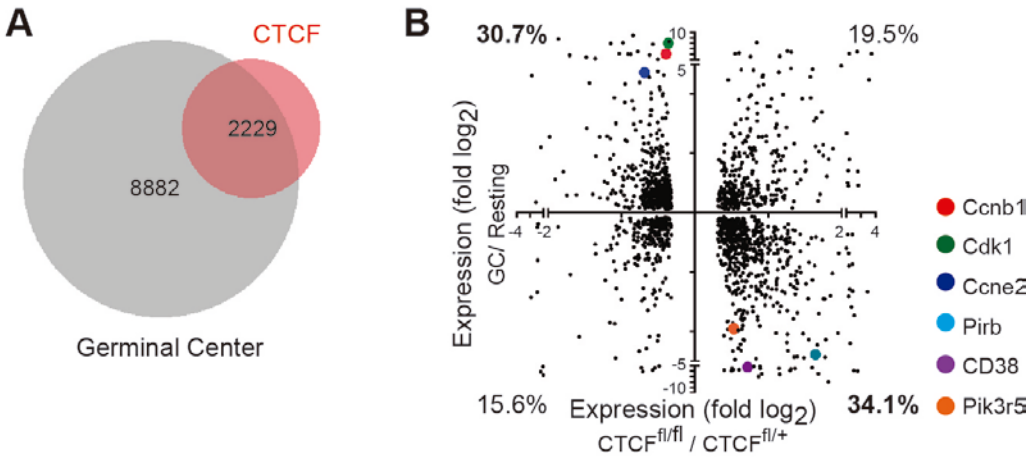

Figure 15. CTCF regulates expression of GC-related genes A) Overlap between genes differentially expressed in GC reaction (RNA-seq from WT PPs GC B cells versus WT resting B cells) and genes differentially expressed upon CTCF depletion in CD3/CD28 T stimulated B cells (RNA-seq from $CTCF^{fl/fl}$ versus $CTCF^{fl/+}$ $hCD2^{+}$ sorted cells). B) Fold change expression profile of differentially expressed genes in GC reaction in vivo compared with resting B cells and in CD3/CD28 stimulated $CTCF^{fl/fl}$ B cells compared with control samples. Percentage of genes in each quadrant is shown.

Pathway enrichment analysis of the genes differentially expressed upon CTCF depletion showed enrichment in terms included in the p53 pathway, B cell receptor signaling, hematopoietic cell lineage and cell cycle (Figure 16), among others.

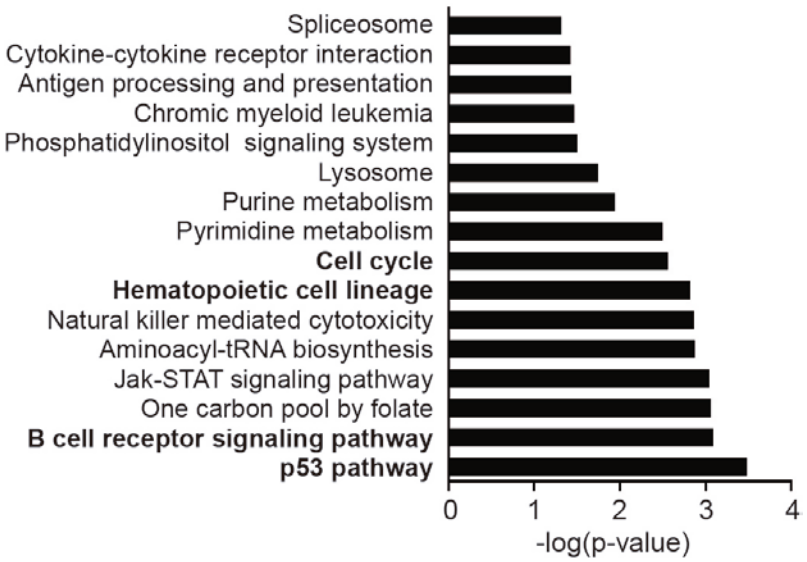

Figure 16. CTCF regulates B cell signaling and cell cycle genes in GC B cells. KEGG pathway enrichment analysis of genes differentially expressed in CD3/CD28 T stimulated $CTCF^{fl/fl}$ versus $CTCF^{fl/+}$ B cells.

Analysis of the individual genes present in each pathway (Figure 17) revealed that CTCF depletion impairs downregulation of Pirb (CD85A), Cr2 (CD21), Fcgr2b (CD32), Pik3r5 (P101) and Ppp3ca (Calcineurin A Alpha), genes, involved in the B cell signaling pathway. On the other hand, CTCF seems to be involved in the upregulation of E2f3, Tfdp1, Ccnb1 (Cyclin B1), Cdc25a, Wee1, Ccne2 (Cyclin E2) and Cdk6, key players of cell cycle progression. This result further highlights the relevance of CTCF as a transcriptional regulator of events critical for the GC reaction.

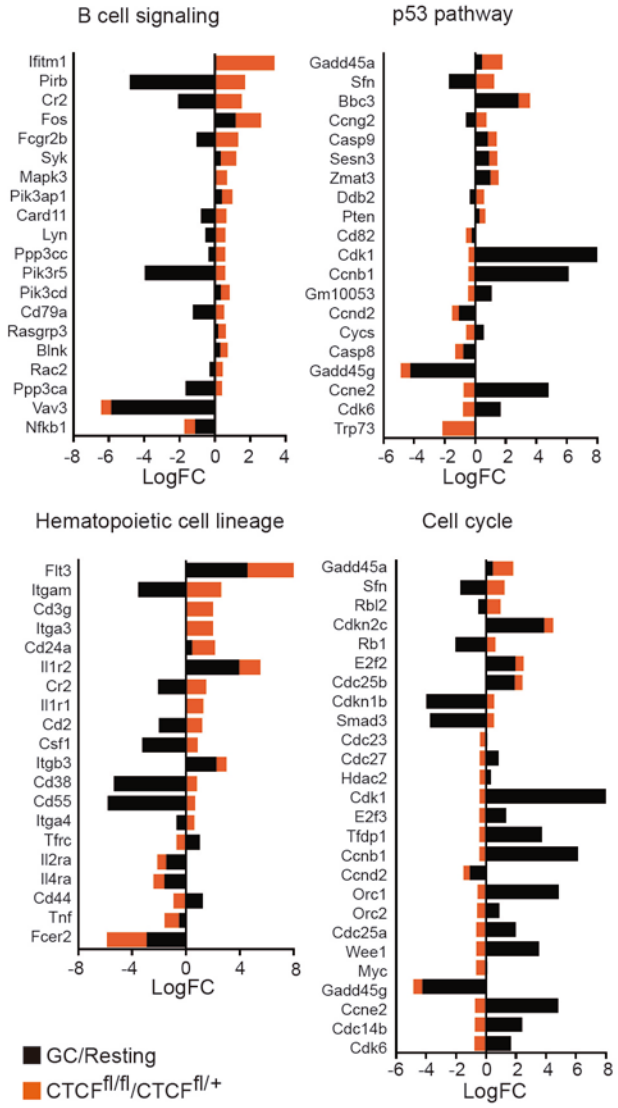

Figure 17. CTCF deficiency alters the expression of genes involved in key pathways in GC B cells. Log fold change (LogFC) representation of the genes included in the KEGG pathways analyzed in Figure 16. Black, GC versus resting B cells fold change. Orange, $CTCF^{fl/fl}$ versus $CTCF^{fl/+}$ CD3/CD28 T stimulated splenic B cells fold change.

1.6. CTCF deficient cells recapitulate key features of plasma cells

One of the main features of the GC B cells is their high proliferation rate (MacLennan, 1994, Allen et al., 2007) required for the generation of large numbers of different Ig variants that undergo affinity-based selection. Our transcriptome results indicate that numerous

genes involved in cell cycle progression are not properly upregulated in CTCF deficient cells. To address whether cell cycle progression was indeed affected by the lack of CTCF we performed Hoechst staining of CD3/CD28 T stimulated $CTCF^{fl/fl}$ and control cells. Indeed, we observed a reduced proportion of S/G2 phase $CTCF^{fl/fl}$ cells (42.8% ± 2.4%) when compared with control cells (52.9% ± 2.4%), as well as an increase in the proportion of G1 phase CTCF deficient cells (55.2% ± 2.5% in $CTCF^{fl/fl}$ cells versus 44.7% ± 2.7% in $CTCF^{fl/f+}$ cells) (Figure 18A, B).

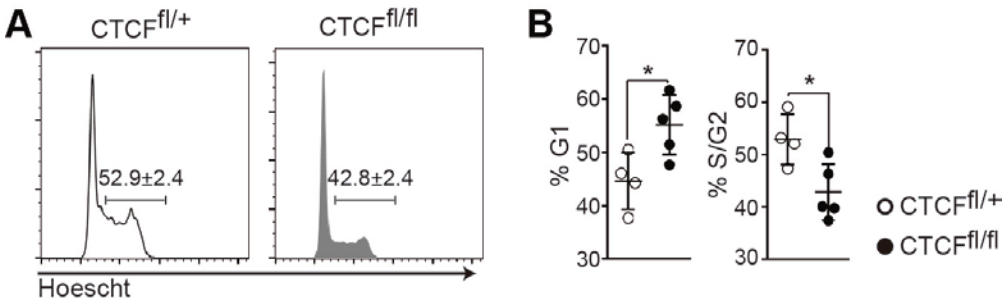

Figure 18. CTCF deficiency impairs cell cycle. A) FACS cell cycle analysis in CD3/CD28 T stimulated $CTCF^{fl/+}$ (n=4) and $CTCF^{fl/fl}$ (n=5) $hCD2^{+}$ cells. Percentages of G1 and S/G2 phases are shown on the right. p(G1)= 0.0242; p(S/G2)= 0.0221. Statistical analysis was done with two-tailed unpaired Student's *t* test.

Transition from naïve to GC B cells involves the activation of a gene-expression program that drives cell proliferation and Ig gene remodeling. In contrast, the transition from GC B cells to antibody secreting PCs involves a transcriptional switch that promotes a halt in cell proliferation to allow cell differentiation. To address whether the maintenance of the GC transcriptional program could be compromised in CTCF deficient cells, we analyzed the shifts in PC transcriptional program (fold change PC/GC (Shi et al., 2015) in the selected genes previously found to be affected by CTCF depletion. In contrast to our data on the naïve to GC transition (Figure 15B) we found a positive correlation between those genes changed at the transition from GC to PC and the ones that are altered upon CTCF depletion.

Specifically, about 60% of the genes analyzed were expressed in the same direction in both conditions (Figure 19A) and about 40% showed anti-correlative expression. Thus, CTCF depletion seems on one hand to counter balance the GC transcriptional program, and on the other, to favor the plasma cell transcriptional program. Among the genes commonly upregulated in CTCF deficient and plasma cells, we found the master transcriptional regulator of the plasma cell differentiation program, Blimp-1 (PRDM1) (Shapiro-Shelef et al., 2003, Shaffer et al., 2002), the plasma cell marker Syndecan-1 and Ig genes, whose expression is greatly increased in antibody secreting cells. Interestingly, a major category of genes whose expression shifts in the GC to PC transition was cell-cycle genes. Analysis of individual genes included in this category showed a sharp downregulation both at the GC to PC transition and in CTCF deficient cells (Figure 19B). These results suggested that CTCF depletion in mature B cells might impair the maintenance of the GC program, promoting the premature acquisition of plasma cells features, such as the upregulation of transcriptional repressor Blimp-1 (confirmed by qRT-PCR, Figure 19C). In support of this notion, we found that Myc and CIITA, two known Blimp-1 targets, were downregulated in CTCF deficient B cells (Figure 19D).

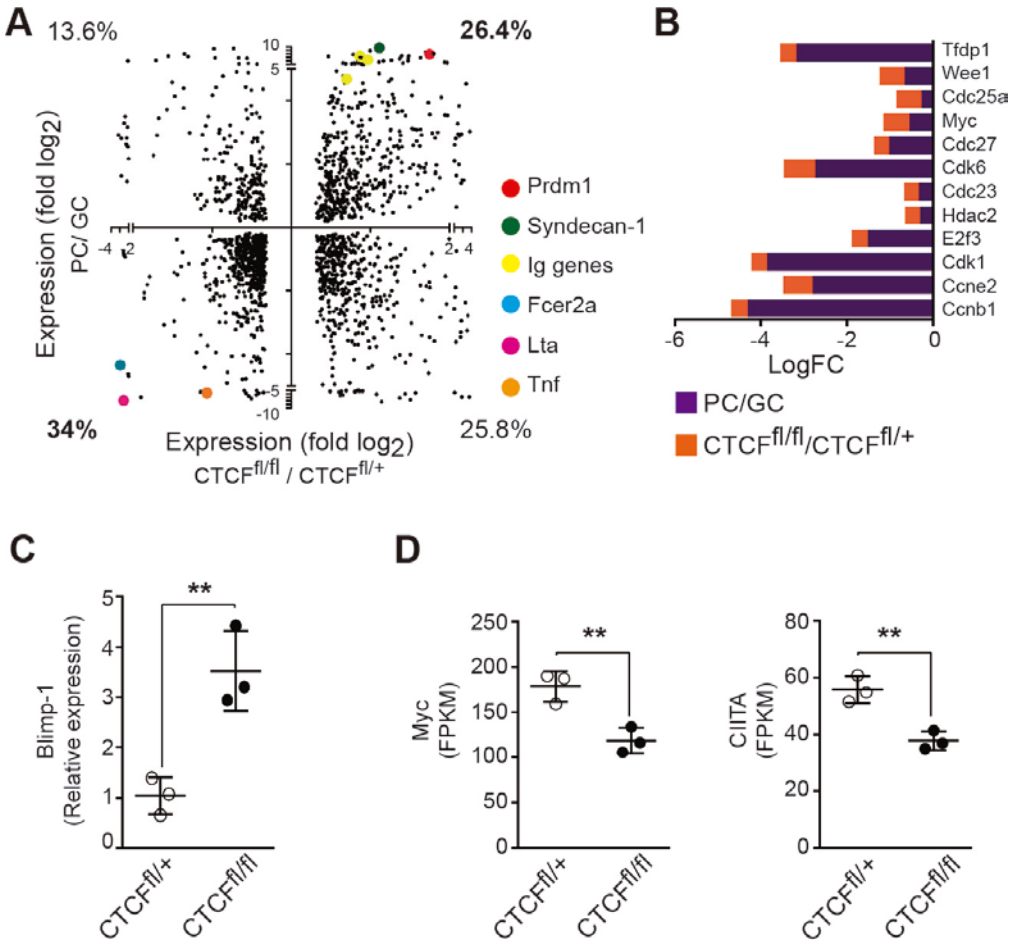

Figure 19. **CTCF deficient cells prematurely express plasma cell genes.** A) Fold change expression profile of differentially expressed genes in PC in vivo compared with GC B cells and in CD3/CD28 T stimulated CTCF$^{fl/fl}$ B cells compared with control samples. Percentage of genes in each quadrant is shown. B) Log fold change (LogFC) representation of cell cycle-related genes. Purple, PC versus GC B cells fold change. Orange, CTCF$^{fl/fl}$ versus CTCF$^{fl/+}$ CD3/CD28 T stimulated splenic B cells fold change. C) qRT-PCR analysis of Blimp-1 expression in CD3/CD28 T stimulated CTCF$^{fl/+}$ and CTCF$^{fl/fl}$ hCD2^{+} cells. p= 0.008. D) RNA-seq analysis of Myc and CIITA expression in CD3/CD28 stimulated CTCF$^{fl/+}$ and CTCF$^{fl/fl}$ hCD2^{+} cells for 48h. p(Myc) = 0.009; p(CIITA)= 0.006. All data are values +/- Standard Deviation (CTCF$^{fl/+}$, white dots; CTCF$^{fl/fl}$; black dots). Statistical analysis was done with two-tailed unpaired Student's t test.

1.7. CD40 signaling restores cell proliferation and Blimp-1 levels in CTCF deficient cells

Blimp-1 is essential for plasma cell differentiation both by extinguishing cell proliferation and by triggering the antibody secretion program. In turn, signaling through the CD40 receptor has been shown to downregulate Blimp-1 levels in B cells (Randall et al., 1998, Knodel et al., 2001, Upadhyay et al., 2014). Thus, we reasoned that if Blimp-1 upregulation were responsible for the proliferation defect found in CTCF deficient cells we could possibly rescue this phenotype by CD40 signaling. We cultured B cells in the presence of T cells stimulated with CD3/CD28 for 24 hours and then added anti-CD40 or left them untreated (Figure 20A). qRT-PCR analysis confirmed once again that CTCF deficient B cells have higher Blimp-1 levels than their littermate controls (Figure 20B). Importantly, this effect was dampened in cells treated with anti-CD40 (Figure 20B) (2.93±0.3 in $CTCF^{fl/fl}$ cells versus 1.40±0.4 in $CTCF^{fl/fl}$+CD40 cells).

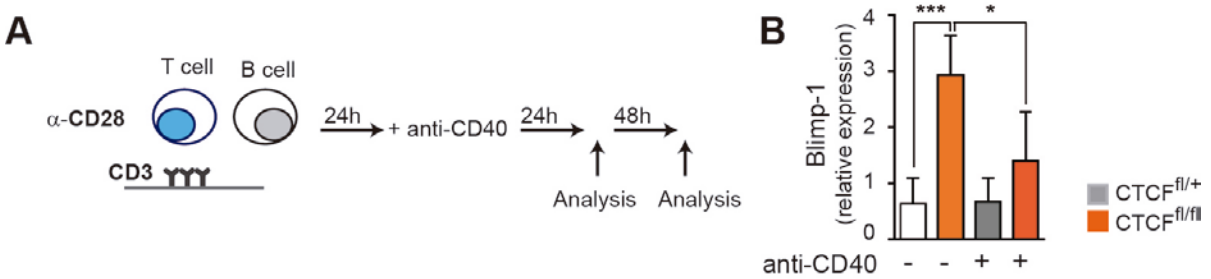

Figure 20. CD40 signaling downregulates Blimp-1 in CTCF deficient cells. $CTCF^{fl/+}$ and $CTCF^{fl/fl}$ splenic B cells stimulated with CD3/CD28 T cells were treated with anti-CD40 antibody. A) Experimental approach. B) Analysis of Blimp-1 mRNA by qRT-PCR in $hCD2^{+}$ $CTCF^{fl/+}$ or $CTCF^{fl/fl}$ cells after 48 hours with or without anti-CD40 treatment. p(***)= 0.0001, p(*)= 0.0159. Mean values +/- Standard Deviation are shown. Statistical analysis was done with two-tailed unpaired Student's t test.

Next, we determined the proportion of B cells in the CD3/CD28 T cells cultures after 48h of treatment with anti-CD40 (Figure 21A, B). We observed that, in the absence of anti-CD40, the proportion of CTCF deficient B cells is reduced when compared with control samples (left

panel) (42.2% ± 4.2% in $CTCF^{fl/+}$ cells versus 22.4% ± 3.1% in $CTCF^{fl/fl}$ cells). However, in the presence of anti-CD40, this difference is attenuated (37.7% ± 4.5% in $CTCF^{fl/fl}$ cells versus 49% ± 7% in $CTCF^{fl/+}$ control cells; p=0.2) (right panel).

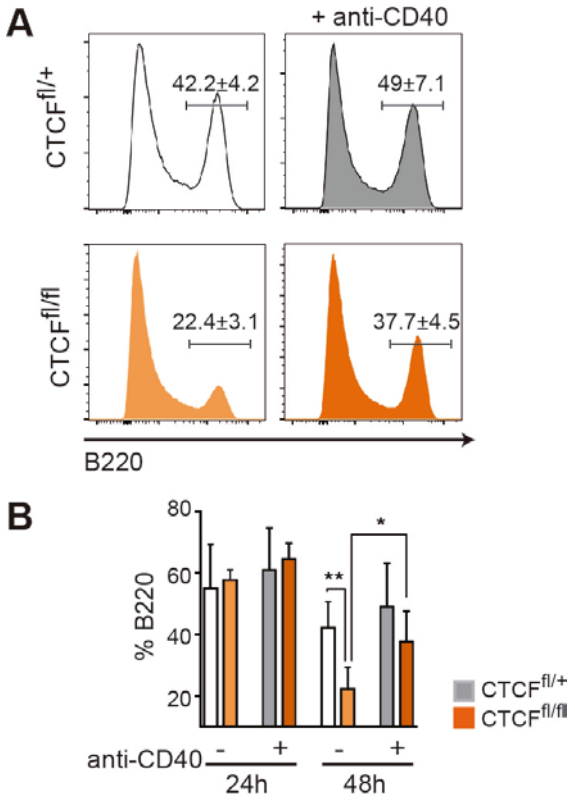

Figure 21. CD40 signaling partially rescues CTCF deficiency. Analysis of B cell proportions in B-T CD3/CD28 co-cultures after 24 or 48 hours of anti-CD40 treatment. A) Representative FACS staining showing B cell population at 48h without (left) or with anti-CD40 treatment (right). B) Quantification of B220 proportion at different time-points after CD40 treatment. p(**)= 0.0063, p(*)= 0.023. Mean values +/- Standard Deviation are shown. Statistical analysis was done with two-tailed unpaired Student's t test.

Finally, we analyzed the cell cycle in these cells (Figure 22A, B). As shown above, we found that CTCF$^{fl/fl}$ cells have a lower proportion of cycling cells than CTCF$^{fl/+}$ cells from littermate controls (35.7%±3.7% versus 21%±2.2%). In contrast, we observed that after treatment with anti-CD40, the proportion of S/G2 cells was not significantly reduced (p=0.07) in CTCF$^{fl/fl}$ (28.5 ± 1) compared with control samples (37 ± 4.3) (right panel). These results indicate that signaling through the CD40 receptor is able to partially rescue the cell cycle impairment associated with CTCF deficiency, presumably through the downregulation of Blimp-1.

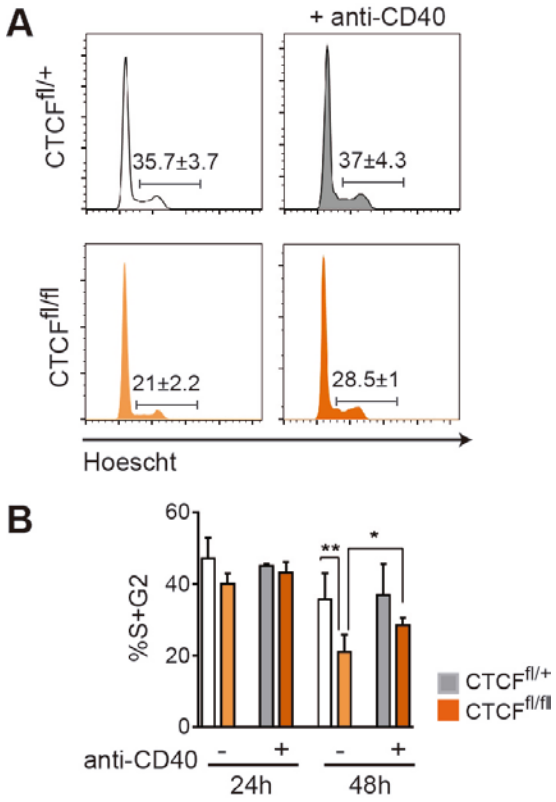

Figure 22. Cell cycle analysis in B-T CD3/CD28 co-cultures after 24 or 48 hours of anti-CD40 treatment. A) Representative FACS staining showing cell cycle at 48h of CD40 treatment. B) Quantification of proliferating cells at different time-points after CD40 treatment. p(**)= 0.0083, p(*)= 0.0126. Mean values +/- Standard Deviation are shown. Statistical analysis was done with two-tailed unpaired Student's t test.

2. Analysis of the contribution of AID to the formation of epithelial neoplasias

AID initiates both SHM and CSR by deaminating cytosines in the variable (SHM) or the switch (CSR) region of the immunoglobulins. However, its activity is not completely restricted to the immunoglobulin genes, and it can promote mutations and DSB followed by illegitimate chromosomal translocations in other regions of the genome.

In addition, in the last years, the expression of AID has been shown to be not restricted to activated B cells, as initially believed. AID expression has been reported in several tissues, including gastric, hepatic and gut epithelia (Endo et al, 2007; Endo et al, 2008; Matsumoto et al, 2007) reviewed in (Marusawa et al, 2011). In these tissues, the expression of AID has been associated with inflammatory contexts and the activation of the NF-κB pathway (Endo et al, 2007) (Matsumoto et al, 2007), and has been claimed to promote the accumulation of mutations in epithelial cells (Matsumoto et al, 2007; Matsumoto et al, 2010; Takai et al, 2009) reviewed in (Marusawa et al, 2011). Given that chronic inflammation in epithelial tissues predisposes to cancer development (Mantovani et al, 2008), the finding that the mutagenic activity of AID can be induced in an inflammatory context has fostered the idea that AID might contribute to or even constitute the link between inflammation and cancer (Takai et al, 2012) reviewed in (Marusawa et al, 2011).

Several gain-of-function mouse models have been generated to address the contribution of AID to neoplastic transformation. Ubiquitous AID overexpression was shown to lead mostly to early T cell neoplasia (Okazaki et al, 2003), among other malignancies. In contrast, B-cell-specific AID overexpression mouse models did not result in lymphomagenesis (Muto et al, 2006; Robbiani et al, 2009) unless the tumor suppressor p53 was removed (Robbiani et al, 2009). Regarding the role of AID in the development of epithelial neoplasias, different studies have made use of ubiquitous AID overexpression mouse models or in vitro assays, but the role of AID expression in epithelial tissues has not been specifically addressed.

2.1. Inflammation-induced AID does not contribute to carcinogenesis

Induction of AID expression in inflammatory conditions has been proposed as a mechanism to explain the link between inflammation and cancer development. As inflammation plays a critical role in the etiology of colorectal and pancreatic ductal adenocarcinoma (reviewed in (Vonderheide and Bayne, 2013, Feagins et al., 2009)), we decided to focus on these two tissues to study the contribution of AID to the formation of neoplasias. To investigate whether inflammatory conditions promote AID expression in these tissues, we treated human epithelial cell lines derived from colorectal adenocarcinoma (LoVo and SW480) and pancreatic adenocarcinoma (AsPC and PaTU) with the pro-inflammatory cytokine TNF-α and measured AID expression by qRT-PCR. TNF-α stimulation increased AID mRNA expression in all cell lines analyzed (Figure 23A, B). To assess whether primary, non-transformed cells were also able to express AID in response to inflammatory stimuli, we generated explants from mouse pancreatic acinar cells and treated them with TNF-α. As with the human tumor cells, mouse primary epithelial cells expressed AID upon exposure to TNF-α (Figure 23C). TNF-α treatment typically induced 4-30 fold increases in AID mRNA levels in the different cell types tested, consistent with previous findings in liver, gastric and colorectal cell lines (Endo et al., 2007, Endo et al., 2008, Matsumoto et al., 2007). Together, these data confirm previous results showing that inflammatory stimuli can trigger AID expression in cell lines originated from human colorectal adenocarcinoma (Endo et al., 2008), and show that pancreatic adenocarcinoma cells and primary pancreatic cells are also responsive to TNF-α treatment.

Inflammation-induced AID expression has been proposed to contribute to or even be the leading cause of some epithelium-derived tumors, such as colorectal adenocarcinoma (Marusawa et al., 2011). To address whether endogenous AID expressed in epithelium under inflammatory conditions could contribute to carcinogenesis, we made use of the well-established model of dextran sulfate sodium (DSS)-induced colitis associated cancer (CAC) (Cooper et al., 2000). $AID^{-/-}$ mice or $AID^{+/-}$ littermates were treated for 10 cycles with 3% DSS and evaluated by pathological criteria. We found that the frequency of oncogenic lesions was not significantly different in $AID^{-/-}$ versus $AID^{+/-}$ mice (Figure 23D). These results

suggested that endogenous AID does not significantly contribute to colorectal adenocarcinoma in the DSS-induced CAC model.

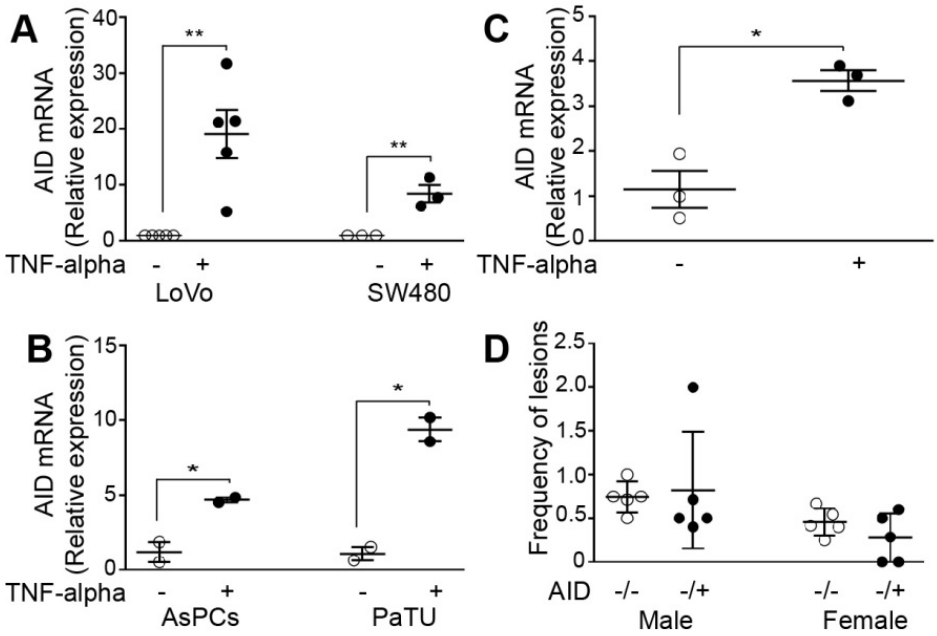

Figure 23. Inflammation-induced AID expression does not contribute to carcinogenesis. (A-C). AID expression was analyzed in colon and pancreatic human cell lines and in pancreas explants from C57BL/6 mice. Samples were treated as indicated with 50ng/ml TNF-a. A) qRT-PCR analysis of AID expression in LoVo and SW480 colon cell lines. n(LoVo)= 5; n(SW480)= 3. p(LoVo)= 0.0017; p(SW480)= 0.0079. B) qRT-PCR analysis of AID expression in AsPC and PaTU-8988S pancreatic cell lines (n=2). p(AsPC)= 0.0369; p(PaTU-8988S)= 0.0119. C) qRT-PCR analysis of AID expression in pancreatic explants from wild-type mice (n=3). P= 0.0242. D) AID-/- or AID+/- mice were treated with 3% DSS for 10 cycles and colonic sections were analyzed by histologic inspection after H/E staining. Graphs represent mean frequency values of adenoma and adenocarcinoma lesions of five independent experiments. n= (AID-/- Males: 28, Females: 35; AID+/- Males: 23, Females: 25). p(Male)= 0.8; p(Female)= 0.246. All data are mean values ± Standard Deviation. Statistical analysis was done with two-tailed unpaired Student's *t* test.

2.2. Generation of conditional AID-expressing mouse models

The absence of a significant contribution of endogenous AID to carcinogenesis in DSS-treated mice could be explained by an insufficient amount of AID in this model. Indeed, AID expression is known to be limiting for its activity in B cells (Sernandez et al., 2008). To directly evaluate whether AID expression can contribute to carcinogenesis, we generated two mouse models for conditional AID expression in epithelial cells of colonic and pancreatic origin (Figure 24A). We introduced an AID-GFP-encoding cassette in the endogenous Rosa26 locus preceded by a transcriptional stop flanked by two loxP sites (R26AID$^{+/KI}$ mice). To achieve specific expression of AID in epithelial cells, we bred R26AID$^{+/KI}$ mice with mice expressing the Cre-recombinase under a villin promoter, which specifically drives expression in colon (el Marjou et al., 2004) (R26AID$^{+/KI}$Villin-CRE$^{+/TG}$ mice), or the pancreas-specific Ptf1 (p48) gene (Kawaguchi et al., 2002) (R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ mice). R26AID$^{+/+}$Villin-CRE$^{+/TG}$ and R26AID$^{+/+}$p48-CRE$^{+/KI}$ mice were used as controls. To confirm that the Rosa26 AID-GFP cassette was functional, we first evaluated the expression of the reporter protein GFP by immunofluorescence in colon of R26AID$^{+/KI}$Villin-CRE$^{+/TG}$ mice and pancreas of R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice (Figure 24B). GFP was expressed in R26AID$^{+/KI}$Villin-CRE$^{+/TG}$ colon and R26AID$^{+/KI}$p48-CRE$^{+/KI}$ pancreas but not in control mice (Figure 24B). We next measured AID transcript levels by qRT-PCR. In R26AID$^{+/KI}$ Villin-CRE$^{+/TG}$ and R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ mice, the amount of AID in the targeted epithelial tissues was similar to that found in B cells activated *in vitro* with LPS+IL4, whereas AID expression in control mice remained at background level (Figure 24C). AID is thus expressed in the epithelium of R26AID$^{+/KI}$ Villin-CRE$^{+/TG}$ colon and R26AID$^{+/KI}$p48-CRE$^{+/KI}$ pancreas at levels known to be functional in B cells.

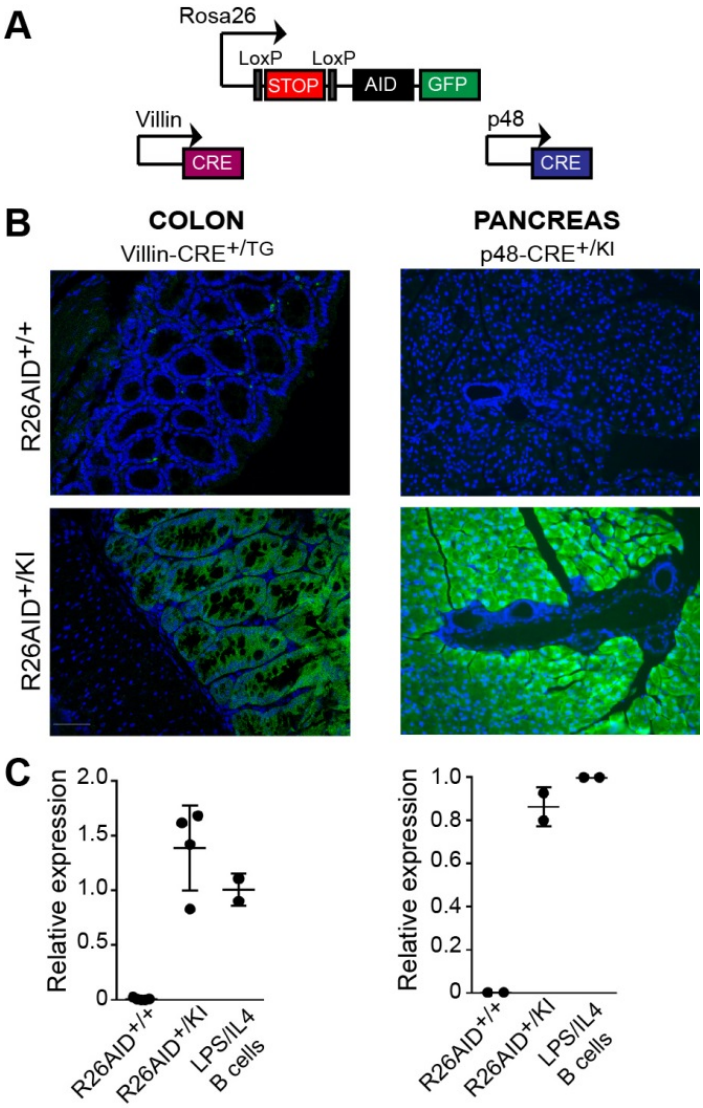

Figure 24. Heterologous AID expression in colon and pancreas. A) Schematic of the constructs used for conditional expression of AID in epithelial cells. An AID-IRES-GFP cassette preceded by a transcriptional STOP flanked by LoxP sites was introduced by homologous recombination within the endogenous Rosa26 locus (R26AID$^{+/KI}$ mice, top). R26AID$^{+/KI}$ mice were bred with Villin-CRE and p48-CRE mice to achieve specific AID expression in colon and pancreas, respectively (bottom). B) GFP immunofluorescence in colonic and pancreatic tissue from R26AID$^{+/KI}$VillinCRE$^{+/TG}$ and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice. Scale bar: 50 μm. C) qRT-PCR analysis of AID expression in colonic and pancreatic tissue from R26AID$^{+/KI}$VillinCRE$^{+/TG}$ and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice. n(R26AID$^{+/+}$ VillinCRE$^{+/TG}$)= 5; n(R26AID$^{+/KI}$VillinCRE$^{+/TG}$)= 4; n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 2, n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 2. LPS+IL4-stimulated B cells are shown as positive control (n=2). Mean values +/- Standard Deviation normalized to LPS+IL4-treated B cells are shown.

2.3. Conditional AID expression in epithelial cells does not promote adenocarcinoma development

To assess the contribution of AID to adenocarcinoma development, we monitored tumor incidence in R26AID$^{+/KI}$ Villin-CRE$^{+/TG}$ and R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ mice. The onset of pancreatic and colorectal adenocarcinoma in a variety of mouse models ranges from 5-6 months to 1-1.5 years (Fodde and Smits, 2001, Aguilar et al., 2004, Martinelli et al., 2016). Therefore, to avoid confounding results arising from spontaneous tumorigenesis in very old mice, we set analysis end-points at 75-100 weeks. Survival of R26AID$^{+/KI}$Villin-CRE$^{+/TG}$ mice was indistinguishable from that of R26AID$^{+/+}$ Villin-CRE$^{+/TG}$ littermate controls (Figure 25A, left). Likewise, survival of R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ did not differ from that of R26AID$^{+/+}$ p48-CRE$^{+/KI}$ controls (Figure 25A, right). To rule out the presence of early malignancies in aged animals, we performed thorough pathological analysis of colon and pancreas sections of all animals, but could not detect any tumor development in R26AID$^{+/KI}$ Villin-CRE$^{+/TG}$ and R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ animals at 75-100 weeks (Figure 25B, C). Expression of AID in colon or pancreatic epithelial cells is thus not sufficient to promote tumor development.

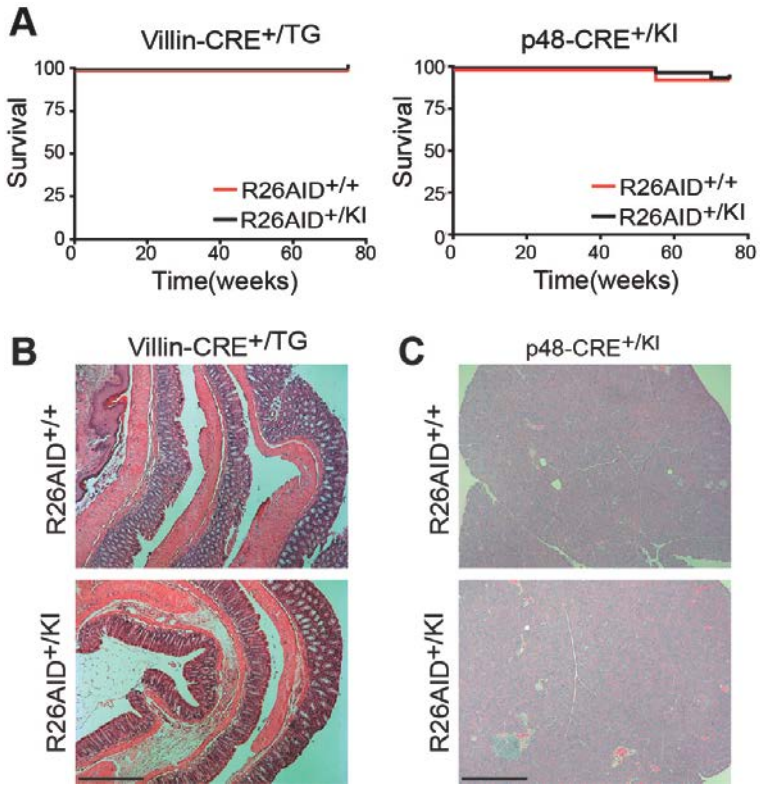

Figure 25. Heterologous AID expression does not promote carcinoma development. A) Kaplan-Meier survival curves for R26AID$^{+/KI}$VillinCRE$^{+/TG}$ (left) (n(R26AID$^{+/+}$ VillinCRE$^{+/TG}$)= 47; n(R26AID$^{+/KI}$VillinCRE$^{+/TG}$)= 38) and R26AID$^{+/KI}$ p48CRE $^{+/KI}$ mice (right) (n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 39; n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 23). B) HE staining in colonic tissue from 75 week-old R26AID$^{+/+}$VillinCRE$^{+/TG}$ (top) and R26AID$^{+/KI}$VillinCRE$^{+/TG}$ (bottom) mice. Scale bar: 500 μm. C) HE staining in pancreatic tissue from 75 week-old R26AID$^{+/+}$p48CRE$^{+/KI}$ (top) and R26AID$^{+/KI}$p48CRE$^{+/KI}$ (bottom) mice. Scale bar: 500 μm.

2.4. AID generates mutations and DNA double strand breaks in pancreatic epithelium

The failure of AID expression to trigger tumorigenesis prompted us to evaluate its activity in epithelial cells. We first analyzed the in vivo mutagenic activity of ectopically expressed AID. The primary target sequences for AID mutagenic activity are immunoglobulin genes; although other genes are known to be susceptible to AID-induced mutagenesis, this occurs at much lower rates (~10^{-4} mutations/bp) and the mechanisms responsible for this

susceptibility are poorly understood. One of the best-characterized requirements for AID activity is that the target sequence be transcriptionally active (Chaudhuri et al., 2003, Ramiro et al., 2003, Pavri and Nussenzweig, 2011). To simplify the mutagenesis analysis we made use of the p48 pancreatic AID expression model to take advantage of the known low complexity transcriptome of acinar cells (MacDonald et al., 2010). AID preferentially targets the consensus hotspots WRCY/RGYW and particularly AGCT motifs (Rogozin and Kolchanov, 1992, Pham et al., 2003, Perez-Duran et al., 2012). Based on this, we analyzed the presence of mutations in 800-900bp downstream of the transcriptional start site of two highly-transcribed genes in pancreas, Elastase1 (Ela1) and Elastase2 (Ela2a), by next generation sequencing, which allows large number of mutations to be analyzed at a very high depth (Perez-Duran et al., 2012). This analysis revealed that mutations are specifically accumulated at the Ela1 and Ela2 genes in R26AID$^{+/KI}$p48-CRE$^{+/KI}$ animals (Figure 26A, B) at frequencies similar to those of other non-Ig genes in B cells (Liu et al., 2008). AID activity was verified for Ela1 by conventional Sanger sequencing (Table 1). In contrast to previous reports (Matsumoto et al., 2007) we did not detect AID-induced mutations at the tumor suppressor gene Trp53, where mutation frequency was identical in R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice and R26AID$^{+/+}$p48-CRE$^{+/KI}$ controls (Table 1).

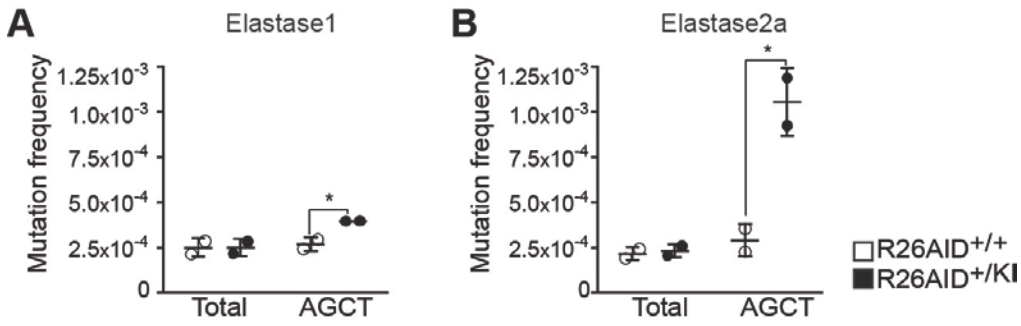

Figure 26. AID expression in pancreas promotes DNA lesions. A,B) Analysis of AID mutagenic activity in Elastase1 (A) and Elastase2a (B) by next generation sequencing. Pancreatic DNA was isolated from pools of R26AID$^{+/KI}$p48CRE$^{+/KI}$ and control R26AID$^{+/+}$p48CRE$^{+/KI}$ 20-week old mice, and then PCR-amplified with specific primers and sequenced as previously described (Perez-Duran et al., 2012). Graphs show cytosine mutation frequency overall (total) or at AGCT hotspots. (n=2).p(Elastase1)= 0.0382; p(Elastase2a)= 0.009. Statistical analysis was done with two-tails unpaired Student's *t* test.

	Genotype	Total clones analyzed	Mutations	Total bp sequenced	Frequency ($x10^{-4}$)
Elastase1	R26AID $^{+/+}$ p48 -CRE$^{+/KI}$	84	4	70018	0.571
	R26AID $^{+/KI}$ p48 -CRE$^{+/KI}$	82	13	69355	1.87
Trp53	R26AID $^{+/+}$ p48 -CRE$^{+/KI}$	66	0	59472	0
	R26AID $^{+/KI}$ p48 -CRE$^{+/KI}$	59	1	53936	0.185

Table 1. Analysis of AID mutagenic activity by Sanger sequencing.

To assess whether AID activity in R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice leads not only to mutations but also to more aggressive lesions, such as DSBs, we quantified γ-H2AX, a histone phosphorylation produced in response to this type of DNA damage. For this analysis, we generated acinar-cell explants from R26AID$^{+/KI}$p48-CRE$^{+/KI}$ and control mice, stained them with anti-γH2AX, and quantified the intensity of staining per nucleus by high-throughput microscopy (HTM). We found that AID expression in R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice promoted a significant increase in the levels of γ-H2AX (Figure 27A, B), indicating that AID generates DSBs in this cellular context.

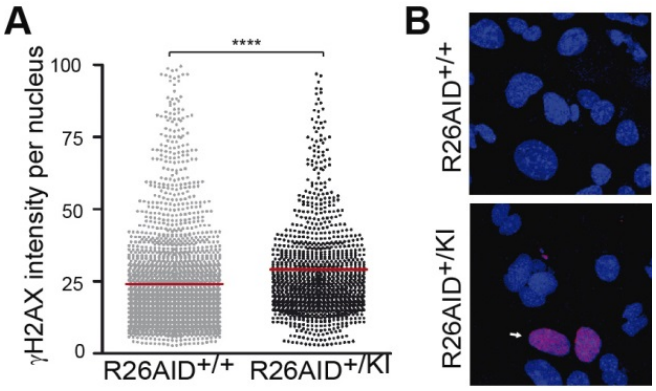

Figure 27. AID expression in pancreas promotes DNA damage. DNA damage was analyzed in pancreas explants from 8-week old mice by staining with anti-γH2AX antibody. A) HTM-mediated quantification of γH2AX intensities per nuclei in pancreas explants cultured in vitro for 6 days. Results of two independent experiments are shown. P<0.0001. B) Representative images of γH2AX staining in pancreas explants from R26AID$^{+/+}$p48CRE$^{+/KI}$ (top) or R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice (bottom) (40X amplification). Statistical analysis was done with two-tailed unpaired Student's *t* test.

2.5. AID induces NKG2D ligands, T cell recruitment and apoptotic cell death in pancreas

Activation of the DNA damage response (DDR) pathway induces the expression of NKG2D ligands in epithelial cells, which are in turn recognized by NKG2D receptors expressed by NK cells and subsets of T cells (Diefenbach et al., 2001, Gasser et al., 2005, Raulet et al., 2013, Champsaur and Lanier, 2010). This crosstalk promotes the elimination of pre-cancerous cells and is therefore a mechanism to prevent tumor development (Guerra et al., 2008). Given that AID expression in pancreas promotes mutations and DNA damage without leading to tumor development, we sought for evidence of precancerous cells and found that pancreas from aged R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice contained more proliferating cells, as assessed by Ki67 staining, than control pancreas (Figure 28A), indicating that pancreatic AID expression leads to an abnormal rate of cell division. The epithelial identity of Ki67^{+} cells was confirmed both by morphology (Fig. 28A, magnified micrograph on the right) and by staining with the epithelium-specific anti-cytokeratin 8 antibody (Figure 28B).

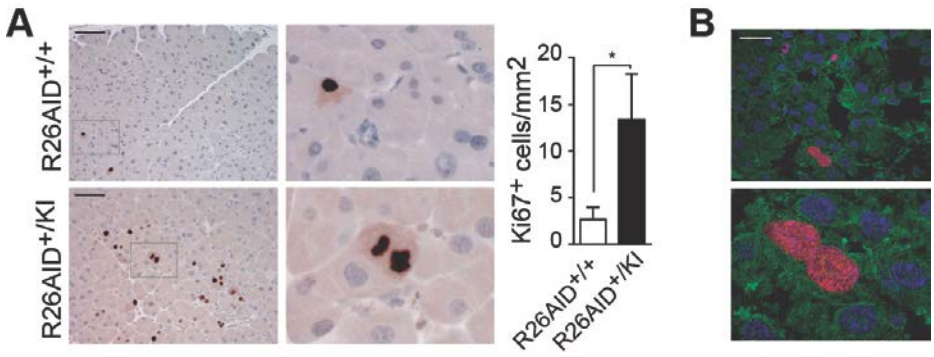

Figure 28. AID expression in pancreas promotes proliferation. Analysis of proliferation was assessed in aged mice by Ki-67 immunohistochemistry. A) Representative images of Ki67 staining in pancreas from aged (75-week-old) R26AID$^{+/+}$p48CRE$^{+/KI}$ (top) or R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice (bottom). Detail is shown on the right. Scale bar: 100 μm. Graph shows quantification of Ki67 positive epithelial cells per mm^2 of tissue (n=8). P= 0.0266. B) Representative immunofluorescence staining of 20-week old R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice: Blue, DAPI; Red, Ki67; Green, CK8. Scale bar: 50 μm. Detail is shown (bottom). Bars show mean values +/- Standard Deviation. Statistical analysis was done with two-tailed unpaired Student's *t* test.

We next asked whether the NKG2D immune surveillance pathway could be in play in R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice. To test this hypothesis, we first analyzed the expression of the Raeε NKG2D ligand in epithelial cells from pancreatic explants of R26AID$^{+/KI}$p48-CRE$^{+/KI}$ and control mice by flow cytometry. Acinar cells from R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice expressed higher levels of Raeε than their control littermates, although the difference was not statistically significant (Figure 29A). To assess whether RAE ligands were expressed by pancreatic cells in vivo, we prepared pancreas extracts from aged (75-week-old) R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice and controls, and measured the amount of five RAE isoforms by droplet digital PCR (ddPCR). With this technique each sample is fractioned into thousands of droplets, in which PCR amplification reactions occur independently, thereby increasing the sensitivity and quantitative potential of the amplification. Amplification of RAE isoforms was detected in more drops from R26AID$^{+/KI}$p48-CRE$^{+/KI}$ samples than from controls (Figure 29B), indicating that AID promotes the expression of NKG2D ligands in pancreas, most likely as a result of DSB and DDR. We found that primary explants from R26AID$^{+/KI}$p48-CRE$^{+/KI}$ tended to be more sensitive to NK-mediated killing than R26AID$^{+/+}$p48-CRE$^{+/KI}$ littermate controls (Figure 29C), indicating that NKG2D ligand expression in AID-expressing pancreas is functional.

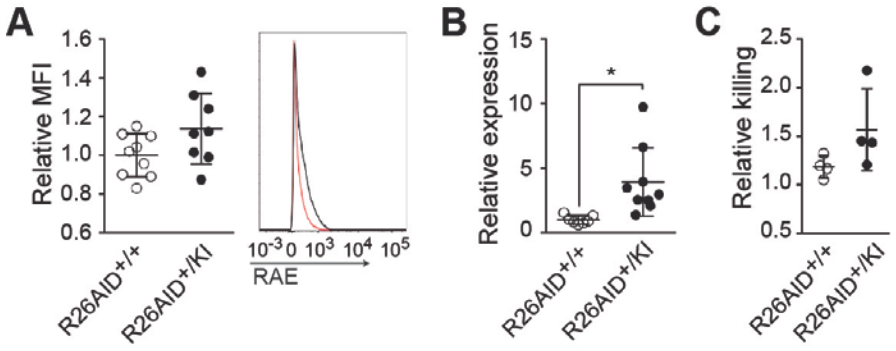

Figure 29. AID expression in pancreas promotes expression of NKG2D ligands. A) Quantitative FACS analysis of RAE expression in pancreatic explants from R26AID$^{+/+}$p48CRE$^{+/KI}$ and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice. Left: Graph shows mean fluorescence intensity. Each dot represents an individual mouse. n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 9; n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 8. P= 0.077. Right: Representative FACS staining for RAE in explants from R26AID$^{+/+}$p48CRE$^{+/KI}$ (red) and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice (black). B) Analysis of RAE expression by ddPCR in aged (75-week-old) mice. Data are presented as the percentage of positive drops normalized to the mean control value. Each point represents an individual mouse and shows the mean amplification from two independent experiments n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 7; n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 9. P= 0.012. C) Analysis of killing activity. Primary explants of pancreatic cells from R26AID$^{+/+}$p48CRE$^{+/KI}$ or R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice were cultured with primary NK cells and killing activity was assessed as described in Methods (n=4). P= 0.19. All data are values +/- Standard Deviation. Statistical analysis was done with two-tailed unpaired Student's *t* test.

We next asked if the expression of RAE ligands promoted the recruitment of immune cells to R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ pancreas in vivo. Hematoxylin-eosin staining of pancreas sections from aged mice clearly revealed the presence of immune infiltrates in AID-expressing pancreas of R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ mice (Figure 30A). The composition of these immune infiltrates was analyzed by antibody staining to detect macrophages (F4/80), B cells (Pax5) and T cells (CD3). The vast majority of cells in the immune infiltrates of R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ mice were CD3^{+} T cells (Figure 30B), with only a negligible contribution from B cells and macrophages (Figure 30C).

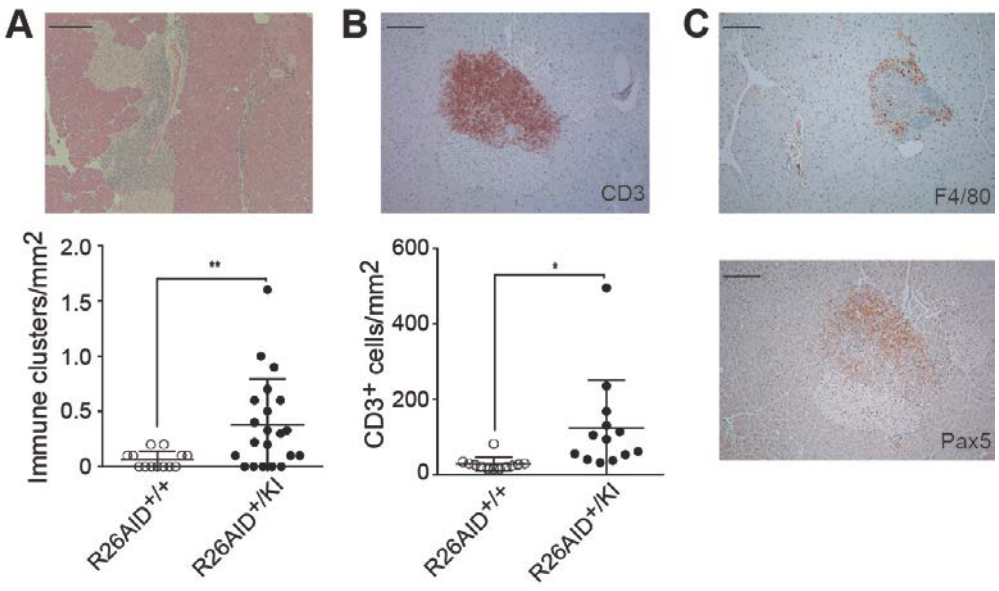

Figure 30. AID expression in pancreas promotes immune cell recruitment. A) Hematoxylin-Eosin (HE) staining of pancreas from aged (75-week-old) R26AID$^{+/+}$p48CRE$^{+/KI}$ and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice. Top: Representative HE staining showing an immune infiltrate in a R26AID$^{+/KI}$p48CRE$^{+/KI}$ mouse. Bottom: Quantification of number of foci per mm^2 of tissue. n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 13; n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 21. P= 0.0098. Scale bar: 200 μm. B) CD3 immunohistochemistry of pancreas from 75-week-old R26AID$^{+/+}$p48CRE$^{+/KI}$ and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice. Top: Representative image of a CD3 infiltrate in a R26AID$^{+/KI}$p48CRE$^{+/KI}$ mouse. Bottom: Quantification of the number of CD3 positive cells per mm2 of tissue. n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 12; n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 13. P= 0.0149. Scale bar: 100 μm. C) Representative image of F4/80 (top) and Pax5 (bottom) staining in a R26AID$^{+/KI}$p48CRE$^{+/KI}$ mouse. All data are values +/- Standard Deviation. Statistical analysis was done with two-tailed unpaired Student's *t* test.

To discount age-related effects, we analyzed 20-week-old mice, finding that the accumulation of T cell infiltrates is detectable in these young animals (Figure 31). The main NKG2D-expressing T-cell subset is the CD8$^+$ population (Raulet et al., 2013), and immunofluorescence analysis of immune infiltrates revealed that a high proportion of the CD3$^+$ infiltrate is composed of CD8$^+$ T cells (Figure 31), a finding consistent with the reported recruitment of CTL cells to pancreatic islets transgenically expressing Raeε (Markiewicz et al., 2012).

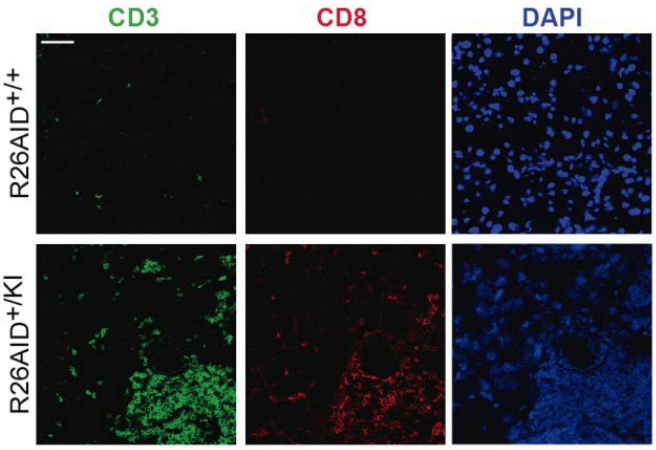

Figure 31. Immune pancreatic infiltrates are mainly composed by CD8$^+$ T cells. Representative immunofluorescence staining of CD3 and CD8 in pancreas from 20-week old R26AID$^{+/+}$p48CRE$^{+/KI}$ and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice. Scale bar: 20 μm.

Finally, we found that aged R26AID$^{+/KI}$p48-CRE$^{+/KI}$ mice had significantly higher levels of pancreatic TNF–α mRNA than control littermates (Figure 32A), indicating that AID promotes the expression of effector cytotoxicity. Consistently, R26AID$^{+/KI}$p48-CRE$^{+/KI}$ pancreas contained cells undergoing apoptotic cell death, detected by caspase3 immunohistochemistry (Figure 32B, p=0.054). Together, these results indicate that heterologous AID expression in pancreas promotes a cytotoxic response, most likely arising from the generation of genotoxic activity and NKG2D ligand expression and the recruitment of NKG2D-expressing CTL cells.

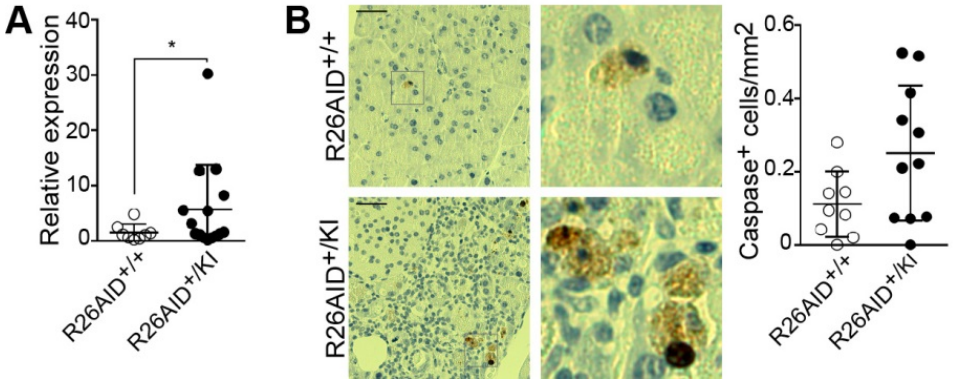

Figure 32. AID expression in pancreas promotes a genotoxic response and cell death. A) TNF-α expression. Total RNA was isolated from pancreas of aged (75-week-old) R26AID$^{+/KI}$p48CRE$^{+/KI}$ and control mice and TNF-α expression was quantified by qRT-PCR. Each dot represents an individual mouse. n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 8; n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 15. P= 0.0392. B) Cell death detection. Pancreas from 75-week old R26AID$^{+/+}$p48CRE$^{+/KI}$ and R26AID$^{+/KI}$p48CRE$^{+/KI}$ mice were stained with anti-caspase-3. Left: Representative staining from an R26AID$^{+/+}$p48CRE$^{+/KI}$ mouse (top) and an R26AID$^{+/KI}$p48CRE$^{+/KI}$ mouse (bottom). Right: Quantification of number of cells per mm^2 of tissue. Each dot represents an individual mouse. n(R26AID$^{+/+}$p48CRE$^{+/KI}$)= 9; n(R26AID$^{+/KI}$p48CRE$^{+/KI}$)= 11. P= 0.0545. Scale bar: 50 μm. All data are values +/- Standard Deviation. Statistical analysis was done with two-tailed unpaired Student's *t* test.

DISCUSSION

1. Analysis of the role of CTCF in the GC reaction

GCs are microstructures where B cells diversify their Ig genes by SHM and CSR and their function is absolutely required for a proper immune response. The establishment of the GC reaction involves a complex transcriptional program with the coordinated expression of gene networks. While the role of master transcriptional regulators of GCs is firmly established, the impact of chromatin structure changes during this reaction is poorly understood. CTCF is believed to play general transcriptional regulatory roles by establishing long-range DNA interactions between distal enhancers and promoters (reviewed in (Ong and Corces, 2014). CTCF activity is essential for embryonic development (Fedoriw et al., 2004).), and in B cells, it has been found to be critical in V(D)J recombination during B cell differentiation in the bone marrow. This defect was associated with a block at the pro-B cell stage. Similarly, in T cells, CTCF deficiency impairs growth and cell proliferation causing cell differentiation arrest (Heath et al., 2008). However the role of CTCF in mature B cells has not previously addressed. We have approached the role of CTCF as a general chromatin remodeler during the GC reaction and have shown that CTCF is absolutely required for the maintenance of the GC reaction in vivo by modulating the expression of genes required for B cell proliferation, including the inhibition of Blimp-1, that programs GC B cells for their entry into PC differentiation (Figure 33). Together, we have described the role of CTCF in orchestrating transcriptional changes required for GC maintenance and PC differentiation.

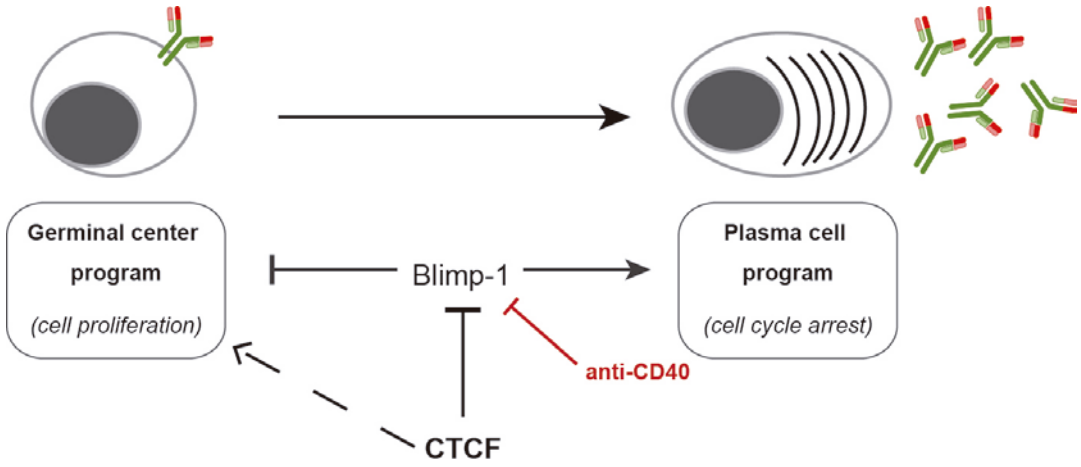

Figure 33 Model of GC B cell regulation by CTCF. CTCF acts upstream of Blimp-1 and different genes controlling the GC program, preventing the premature transition to PC. In this context, anti-CD40 stimulation partially rescues CTCF associated phenotype by downregulating Blimp-1.

1.1. Culture conditions

Our results show that susceptibility to CTCF deficiency depends on the context on which B cells have been activated. In vivo, mature B cell activation in the periphery is triggered by the contact with exogenous antigen and T cells that in turn have been stimulated by the same antigen. These interactions initiate a cascade that allows B cell proliferation and the production of high-affinity and switched antibodies. In vitro, B cell activation can be mimicked by the addition of particular cytokines, together with mitogens (such as LPS) that induce both CSR and cell proliferation (reviewed in (Stavnezer et al., 2008)). However, it has long been known that these conditions fail to fully recapitulate in vivo GC reaction; for instance, they do not trigger SHM in the variable region of Ig genes (Manser, 1987, Di Noia and Neuberger, 2007). Our comparative transcriptome analysis across different B cell stimulation landscapes has revealed that despite sharing a core of a few thousands of common genes, a considerable fraction of transcriptional shifts taking place in GC B cells are not mirrored by LPS/IL4 stimulation. Interestingly, we show that stimulation of B cells in vitro in the presence of T cells activated with CD3 and CD28 is a

better surrogate of in vivo stimulation in GCs, as revealed by whole transcriptome analysis. We believe that this finding will have an important impact by allowing the study of B cell biology in vitro in a more physiological manner.

1.2. B cells susceptibility to CTCF deficiency

Our results shown that susceptibility to CTCF deficiency depends on the context on which B cell have been activated. *In vivo,* mature B cell activation in the periphery is triggered by the contact with exogenous antigens and antigen-related T cells. These interactions initiate a cascade that allows B cell proliferation and the production of high-affinity and switched antibodies. *In vitro*, B cell activation can be mimicked by the addition of particular cytokines, together with mitogens (such as LPS) that induce both CSR and cell proliferation (reviewed in (Stavnezer et al., 2008)). However, despite the efficiency in inducing CSR in vitro, current protocols are known to do not completely recapitulate *in vivo* GC reaction, for example, by the absence of SHM in the variable region (Manser, 1987). We have performed a transcriptional analysis comparing the different programs activated in each condition. Despite both conditions share a core of common genes, expected considering that we compared activated versus resting B cells, LPS/IL4 stimulation does not fully recapitulates GC activation at transcriptional level, in key pathways of GC response, such as cell cycle. We have improved in vitro B cell stimulation by using T cells to stimulate B cells, better mimicking in vivo GC environment at transcriptional level and reflected in CTCF deficient associated phenotype. This finding might be relevant to approach B cell biology in vitro in a more physiological manner.

Regardless of the improvement in mimicking *in vivo* conditions, we were not able to fully reproduce *in vitro* the dramatic effect observed associated to CTCF deficiency *in vivo*. One possible explanation to this difference might lie in the differences of cell proliferation rate. GC B cells present one of the fastest rates of cell division, completing a cell cycle between 6-12 hours (Hauser et al., 2007, Allen et al., 2007). We have seen that most of the phenotype associated to CTCF deficiency might be ascribed to a defective cell cycle.

However, in vitro, the proliferation rate observed does not reach the in vivo one. That might explain why the effect of CTCF deficiency *in vitro* is not as prominent as *in vivo*. Another possibility to consider is the kinetics of CTCF depletion. Messenger or protein stability could be different in both systems, presumably smaller in vivo, what could explain, together with differences in cell proliferation, the magnification of the phenotype associated to CTCF deficiency.

1.3. CTCF control of transcription in GC

We have found that CTCF controls transcription of more than 2000 genes in CD3/CD28 stimulated B cells, mediating both repression and activation. This result contrasted with that found in LPS/IL4 stimulated cells where 50 genes were differentially expressed. We consider that transcriptional differences observed in our two *in vitro* systems might not be due to differences in CTCF binding between them, given the high conservation of CTCF binding among different species (Schmidt et al., 2012) as well as among different cell types (Wang et al., 2012, Kunarso et al., 2010, Cuddapah et al., 2009, Heintzman et al., 2009). Otherwise, we consider that the presence of exogenously added cytokines activates survival and proliferation pathways that promote B cell survival independently on CTCF.

High-throughput studies have revealed that the pattern of CTCF binding in the genome is similar to transcription factors such as TAF1 and its binding is enriched in regions where gene density is increased, suggesting a role of CTCF in transcription control (Kim et al., 2007). However, it binds far from promoters, in intergenic regions, making difficult to determine the direct role of CTCF in transcription control (Kim et al., 2007). A recent study has provided evidences that CTCF generates topological foci that allow regulation of constitutive genes. Within these domains, interaction with RNApolII will determine transcription of cell-type specific genes (Zhonghui, T. Cell 2015). However, direct role of CTCF in global transcription has not been addressed. Here, we have performed an extensive analysis of the impact at transcriptional level of CTCF depletion in activated B cells, showing

that CTCF controls gene networks involved in key aspects of B cell biology, modulating both its activation and its repression.

Our transciptome results have pointed the relevance of CTCF in controlling key pathways of GCs such as B cell receptor signaling and cell cycle. Although we could demonstrate the impairment in cell cycle upon CTCF depletion, in this system, addressing the role of B cell receptor remains technically difficult. Analysis of BCR signaling requires short-time stimulation of B cells with anti-IgM. However, we cannot stimulate in this way our naïve $CTCF^{fl/fl}$-$AIDCRE^{TG/+}$ cells, thus CTCF depletion is only produced after AID expression.

Considering the relevance of these pathways in B cell activation biology, together with the complete absence of CTCF GC deficient cells in vivo, we could consider that in vivo, GC is not initiated. However, the conditional model used in this work, as previously mentioned, implies GC program activation before CTCF elimination. Thus, we can affirm that CTCF is required for the GC program maintenance in this context. The involvement of CTCF in GC initiation should be addressed in mature CTCF B cells, obtained, for example, using a CRE-recombinase under CD19 promoter.

1.4. CTCF and transition to PC

The diminished proliferation of CTCF deficient cells together with the upregulation of key PC factors, such as Blimp-1 or Syndecan1, suggested that CTCF maintains the GC program by repressing the transition to PCs. Upregulation of Blimp-1 upon CTCF depletion has been previously found in a lymphoma cell line (Batlle-Lopez et al., 2015). Our results expand this observation to primary B cells. In addition, our transcriptome analysis comparing CTCF deficient B cells with PC showed that not only Blimp-1 and Syndecan1, but a number of genes were expressed in the same direction. Our results strongly suggest a role of CTCF in balancing decision between GC and PC program.

Blimp-1 is considered the master regulator of PC differentiation (Shaffer et al., 2002, Shapiro-Shelef et al., 2003) by its ability to modulate gene networks involved in cell cycle and antibody secretion. In our system, we have seen direct Blimp-1 targets downregulated, such as CIITA and Myc that could, at least partially, explain the phenotype associated to CTCF deficiency. However, we cannot rule out additional effects directly dependent on CTCF. In this regard, in T cells, CTCF depletion has been associated to a strong upregulation of the negative regulators of cell cycle p21 (CDKN1A) and p27 (CDKN1B) (Heath et al., 2008), what involves CTCF in directly cell cycle control.

1.5. Rescue of CTCF-deficient associated phenotype by CD40 signaling

We have found that anti-CD40 treatment of $CTCF^{fl/fl}$ cells partially rescues the phenotype associated to CTCF deficiency. CD40 engagement triggers a number of pathways in B cells, among cell proliferation, survival and switching can be mentioned (reviewed in (Elgueta et al., 2009)). In addition, we and others have shown that signaling through CD40 is able to down regulate Blimp-1. This down regulation has been postulated to maintain the GC program, avoiding PC differentiation (Randall et al., 1998, Knodel et al., 2001). In our model, we observed that CD40 treatment recovered Blimp-1 levels in CTCF deficient cells, increasing cell proliferation and survival. These results would be in accordance with previous reports in which arrest of cell cycle by Blimp-1 have been shown (Shaffer et al., 2002). Although we cannot rule out the possibility that anti-CD40 treatment increases B cell survival and proliferation independently on Blimp-1, the effect in survival and proliferation was specifically observed in CTCF deficient, but not in control samples.

Together, these results uncover a new role for CTCF in the maintenance of the GC program, by promoting cell proliferation and by inhibiting PC differentiation through Blimp-1 inhibition.

2. Contribution of AID to the generation of neoplasias

2.1. AID expression under inflammatory conditions

AID expression has been shown to be induced in epithelium under inflammatory conditions. *In vitro*, several studies have demonstrated that AID is induced by TNF-α treatment in colon, liver and cholangiocarcinoma-derived cell lines (Marusawa et al., 2011). In this regard, our analyses of AID expression in response to TNF-α have confirmed previous data in colonic cell lines and have expanded these observations to pancreatic cell lines. In addition, we found that primary pancreatic epithelium is also sensitive to TNF-α, suggesting that AID expression can indeed take place in non-transformed cells in pro-inflammatory contexts.

However, despite these *in vitro* evidences of AID induction after inflammatory insults, *in vivo*, only correlative results between the presence of inflammation-associated carcinomas and the expression of AID have been provided (Komori et al., 2008, Kou et al., 2007, Matsumoto et al., 2007). We wanted to further explore the physiological relevance of AID expression in promoting epithelium malignant transformation. We found that AID deficiency does not reduce the incidence of oncogenic lesions in an inflammation-induced carcinoma model in colon. Our results contrast with the finding that AID deficiency reduces colon carcinogenesis in a model of chronic inflammation, triggered by IL10 deficiency (Takai et al., 2012). It may be that the effect reported by Takai et al. is not specifically driven by epithelial cells, but rather by B cells in $IL10^{-/-}$ mice, a possibility that could be tested using conditionally rather than constitutively AID-deficient animals.

2.2. Models of AID expression in epithelium

Previous reports have claimed that AID heterologous expression leads to epithelial cell neoplasia in various tissues (Okazaki et al., 2003, Takai et al., 2009). In those studies overexpression was achieved with transgenes, none of which was epithelium specific. In the

model generated by Okazaki and colleagues, AID expression was achieved in all the tissues by the strong CAG promoter. However, it remains unclear the levels of AID expression reached in these tissues, by the lack of a quantitative measure of AID expression. On the other hand, Takai and colleagues presented a conditional AID expression transgenic model (Muto et al., 2006) in which the expression of AID was achieved by the tissue-nonspecific alkaline phosphatase promoter (TNAP), which is a marker of primordial germ cells. In this work, although AID expression and protein are shown, the levels obtained in the different tissues upon transgene expression are not compared with the physiological levels found in an activated B cell. In contrast to these studies, here we have developed epithelium-specific conditional knock-in model, directing AID expression to colon and pancreas epithelium, and avoiding both widespread expression and transgene-derived artifacts. Remarkably, our models allowed B-cell-like AID expression levels, a basic prerequisite for AID to be competent, given that its levels are rate limiting in its native context (Sernandez et al., 2008, Takizawa et al., 2008).

Regarding the contribution of AID to carcinoma development, results differ among the different models presented. In Okazaki's work, survival of the five transgenic lines generated varied between 40-90 weeks. The authors found a huge proportion of enlarged spleen and lymph nodes, with a high prevalence of T cell lymphomas, as well as lung micro-adenomas. Stickling, the type of tumor and the severity of the phenotype varied among the different transgenic lines generated, as well as with ongoing mouse generations. Although the authors ascribed this effect to the accumulation of AID-derived mutation in the germline, it could be also explained by effects derived from the transgene insertion by itself. On the other hand, in the model presented by Takai, survival of AID-Tg mice was similar to that found in control mice. In this case, authors found a low prevalence of hepatocellular carcinoma (HCC) (27% of mice at 90 weeks), as well as lung and stomach cancer (6.7% of mice at 90 weeks). In contrast to these studies, our models of AID expression in pancreas and colon did not shown any evidence of carcinoma development. Discrepancy found with previous reports might be explained by the used of knock-in models, and by the specific expression of AID in epithelial cells. Our results suggest that AID expression is not enough to promote tumor development.

2.3. AID activity in epithelial cells

AID preferentially mutates specific regions of the Ig genes during the processes of SHM and CSR. However, in the recent years has become clear the ability of AID of mutating non-Ig genes (Liu et al., 2008). Addressing the mutational activity of AID out of Ig genes is difficult given the relative low frequency (10^{-5} to 10^{-3} mutations per base pair) and the widespread nature of AID mutations. Traditionally, AID activity has been measured by Sanger-sequencing at specific sites of interest. Here, we have applied a PCR-based NGS approach that allows the detection of mutations at frequencies bellow 10^{-4} per base pair (Perez-Duran). Transcription is the main requirement for AID activity. For this reason, we focused our mutational analysis on pancreas, given the low complexity transcriptome of acinar cells (MacDonald et al., 2010). AID mutagenic activity was detected in two highly expressed pancreatic genes (Ela1 and Ela2), which is expected from the well-established link between AID activity and transcription of its target genes (Ramiro et al., 2003, Chaudhuri et al., 2003, Pavri and Nussenzweig, 2011, MacDonald et al., 2010). Contrary to previous reports (Takai et al., 2009, Endo et al., 2008, Matsumoto et al., 2007), we did not find evidences of AID-induced mutations at the tumor suppressor Trp53. This discrepancy might reflect differences in transcriptional programs activated in different cell types, determining AID accessibility to target sequences.

In our model of AID expression in pancreas, we found that AID expressed ectopically in pancreatic cells is able not only to mutate non-immunoglobulin genes but also to generate genotoxic DSBs, suggesting that in this context there are no obvious mechanisms for negative regulation of AID activity, as they were previously proposed for transgenic AID expression in B cells (Muto et al., 2006). Repair of AID-induced mutations promotes generation of abasic sites on DNA that result in a single-strand DNA break (SSB). Generation of DSBs by AID activity is supposed to occur when two SSBs sites appear in close proximity in opposite strands, by the action of different repair pathways that include the action of exo- and endonucleases. In switch regions, DSBs are supposed to be favor because of the high concentration of AID hotspots in both strands. However, in epithelial cells, generation of DSBs by AID activity has not been addressed. In our model, we have detected a significant

increase in the proportion of DSBs generated by AID activity in epithelium. In this context, DSBs generation cannot be explained by the accumulation of several AID hotspots in a particular disposition, as is found in the switch regions, suggesting that the mechanisms involved in repairing AID damage in this context might promote the generation of DSBs.

2.4. Epithelium protection against AID activity

Data presented in this work indicate that AID activity in pancreas does not promote pancreatic carcinogenesis. We propose that an immune surveillance process might be activated by AID activity in pancreas, blocking progression of transformed cells. Upregulation of NKG2D ligands in cancerous or stressed tissues has been reported to tag cells to be eliminated by the immune system (Gasser and Raulet, 2006). Here, we show that AID expression triggers the upregulation of Rae in pancreas, in line with the published finding that AID promotes an NKG2D immune response in B cells infected with the Abelson murine leukemia virus (Gourzi et al., 2006). We have found that NKG2D-expressing pancreatic cells mainly recruit $CD8^+$ cells, in accordance with previous reports showing the role of $CD8^+$ T lymphocytes in the elimination of transformed cells (Raulet et al., 2013). Likewise, we have found in these tissues an increased cell death, suggesting a genotoxic response promoted by the immune infiltration found in epithelium (Model depicted in figure 34). The lymphomagenic potential of AID was previously shown to be dampened in B cells (Robbiani et al., 2009, Muto et al., 2006), where p53 exerts a cell-intrinsic tumor suppressor function (Robbiani et al., 2009). Here we provide evidence of a further protective mechanism triggered by AID activity and carried out through an extrinsic immune surveillance pathway in epithelial tissues.

Together, data presented in this part of the work uncover a new protective mechanism counter-acting AID oncogenic potential through an extrinsic immunosurveillance pathway in epithelial cells. These findings challenge the view that AID contribute significantly to inflammation-related carcinogenesis.

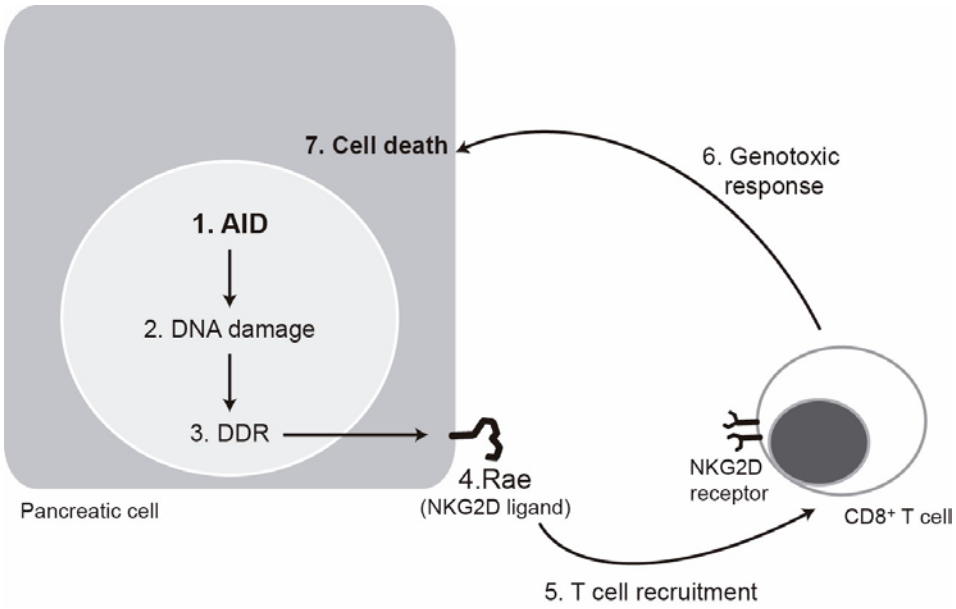

Figure 34. AID-expressing epithelia are protected from neoplastic transformation by an NKG2D surveillance pathway. AID expression in pancreas promotes a response that involves NKG2D ligand expression, $CD8^+$ T cell recruitment and epithelial cell death.

CONCLUSIONES

1. CTCF es un remodelador de la cromatina absolutamente necesario para la reacción de centro germinal, bien sean constitutivas o *de novo*.

2. El contexto en el cual CTCF es eliminador de las células B determina la susceptibilidad de dichas células a la falta de este factor.

3. A nivel transcripcional, la estimulación mediante LPS/IL4 difiere con respecto a la reacción de GC in vivo en rutas clave para la biología de estas células. La estimulación de células B *in vitro* a través de la activación de células T, recapitula mejor el transcriptoma de los GCs in vivo.

4. CTCF controla a nivel transcripcional rutas clave de las células B de GC como la señalización a través del BCR o el ciclo celular.

5. La deficiencia de CTCF altera el mantenimiento del programa de GC y provoca la transición prematura hacia el programa de PC.

6. El silenciamiento del regulador maestro de la diferenciación hacia PC Blimp-1, rescata de forma parcial el fenotipo asociado a la deficiencia de CTCF en células B.

7. La inducción de inflamación promueve la expresión de AID tanto en líneas como en células primarias epiteliales.

8. Los modelos R26AID$^{+/KI}$ Villin-CRE$^{+/TG}$ y R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ permiten la expresión específica de AID en colon y páncreas, respectivamente, a niveles similares a los obtenidos en células B activadas.

9. La expresión de AID en colon y páncreas no es suficiente para promover el desarrollo tumoral.

10. La expresión de AID en páncreas promueve una respuesta citotóxica, resultante del acúmulo de daño en el DNA, la expresión de ligandos NKG2D y el reclutamiento de linfocitos T CD8^{+}.

CONCLUSIONS

1. CTCF is a chromatin remodeler absolutely required for constitutive and de novo GC reactions in vivo.

2. Stimulation context in which CTCF is depleted in B cells determines the susceptibility of these cells to the lack of this factor.

3. Transcriptionally, LPS/IL4 in vitro stimulation differs to in vivo GC cells in key pathways of GC B cell biology. In vitro stimulation of B cells through T cell activation better mimics in vivo context.

4. CTCF controls at transcriptional level key pathways of GC B cells such as B cell receptor pathway and cell cycle.

5. CTCF deficiency impairs the maintenance of the GC program and determines an imbalance through the PC program.

6. Downregulation of the master regulator of PC differentiation Blimp-1 partially rescues the phenotype associated to CTCF deficiency in B cells.

7. Inflammatory stimulation triggers AID expression in epithelial cell lines and in non-transformed primary pancreatic cells.

8. R26AID$^{+/KI}$ Villin-CRE$^{+/TG}$ and R26AID$^{+/KI}$ p48-CRE$^{+/KI}$ models allow specific expression of AID in colon and pancreas, respectively, at levels similar to those found in activated B cells.

9. AID expression in colon and pancreas is not sufficient to promote tumor development.

10. AID expression in pancreas promotes a cytotoxic response, most likely arising from the generation of genotoxic activity and NKG2D ligand expression and the recruitment of CD8^{+} T cells

BIBLIOGRAPHY

AGUILAR, S., COROMINAS, J. M., MALATS, N., PEREIRA, J. A., DUFRESNE, M., REAL, F. X. & NAVARRO, P. 2004. Tissue plasminogen activator in murine exocrine pancreas cancer: selective expression in ductal tumors and contribution to cancer progression. *Am J Pathol,* 165**,** 1129-39.

ALLEN, C. D., OKADA, T., TANG, H. L. & CYSTER, J. G. 2007. Imaging of germinal center selection events during affinity maturation. *Science,* 315**,** 528-31.

AOUFOUCHI, S., FAILI, A., ZOBER, C., D'ORLANDO, O., WELLER, S., WEILL, J. C. & REYNAUD, C. A. 2008. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med,* 205**,** 1357-68.

BACHL, J., CARLSON, C., GRAY-SCHOPFER, V., DESSING, M. & OLSSON, C. 2001. Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J Immunol,* 166**,** 5051-7.

BASSING, C. H., SWAT, W. & ALT, F. W. 2002. The mechanism and regulation of chromosomal V(D)J recombination. *Cell,* 109 Suppl**,** S45-55.

BASSO, K. & DALLA-FAVERA, R. 2010. BCL6: master regulator of the germinal center reaction and key oncogene in B cell lymphomagenesis. *Adv Immunol,* 105**,** 193-210.

BASU, U., CHAUDHURI, J., ALPERT, C., DUTT, S., RANGANATH, S., LI, G., SCHRUM, J. P., MANIS, J. P. & ALT, F. W. 2005. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature,* 438**,** 508-11.

BATLLE-LOPEZ, A., CORTIGUERA, M. G., ROSA-GARRIDO, M., BLANCO, R., DEL CERRO, E., TORRANO, V., WAGNER, S. D. & DELGADO, M. D. 2015. Novel CTCF binding at a site in exon1A of BCL6 is associated with active histone marks and a transcriptionally active locus. *Oncogene,* 34**,** 246-56.

BRANSTEITTER, R., PHAM, P., CALABRESE, P. & GOODMAN, M. F. 2004. Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase. *J Biol Chem,* 279**,** 51612-21.

BRANSTEITTER, R., PHAM, P., SCHARFF, M. D. & GOODMAN, M. F. 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci U S A,* 100**,** 4102-7.

BRAR, S. S., WATSON, M. & DIAZ, M. 2004. Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J Biol Chem,* 279**,** 26395-401.

CAGANOVA, M., CARRISI, C., VARANO, G., MAINOLDI, F., ZANARDI, F., GERMAIN, P. L., GEORGE, L., ALBERGHINI, F., FERRARINI, L., TALUKDER, A. K., PONZONI, M., TESTA, G., NOJIMA, T., DOGLIONI, C., KITAMURA, D., TOELLNER, K. M., SU, I. H. & CASOLA, S. 2013. Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. *J Clin Invest,* 123**,** 5009-22.

CALADO, D. P., SASAKI, Y., GODINHO, S. A., PELLERIN, A., KOCHERT, K., SLECKMAN, B. P., DE ALBORAN, I. M., JANZ, M., RODIG, S. & RAJEWSKY, K. 2012. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol,* 13**,** 1092-100.

COOPER, H. S., MURTHY, S., KIDO, K., YOSHITAKE, H. & FLANIGAN, A. 2000. Dysplasia and cancer in the dextran sulfate sodium mouse colitis model. Relevance to colitis-associated neoplasia in the human: a study of histopathology, B-catenin and p53 expression and the role of inflammation. *Carcinogenesis,* 21**,** 757-68.

CRISCUOLO, A. & BRISSE, S. 2014. AlienTrimmer removes adapter oligonucleotides with high sensitivity in short-insert paired-end reads. Commentary on Turner (2014) Assessment of insert sizes and adapter content in FASTQ data from NexteraXT libraries. *Front Genet,* 5**,** 130.

CUDDAPAH, S., JOTHI, R., SCHONES, D. E., ROH, T. Y., CUI, K. & ZHAO, K. 2009. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res,* 19**,** 24-32.

CHAMPSAUR, M. & LANIER, L. L. 2010. Effect of NKG2D ligand expression on host immune responses. *Immunol Rev,* 235**,** 267-85.

CHAUDHURI, J. & ALT, F. W. 2004. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol,* 4**,** 541-52.

CHAUDHURI, J., TIAN, M., KHUONG, C., CHUA, K., PINAUD, E. & ALT, F. W. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature,* 422**,** 726-30.

CHEN, H., TIAN, Y., SHU, W., BO, X. & WANG, S. 2012. Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PLoS One,* 7**,** e41374.

DE YEBENES, V. G., BELVER, L., PISANO, D. G., GONZALEZ, S., VILLASANTE, A., CROCE, C., HE, L. & RAMIRO, A. R. 2008. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J Exp Med,* 205**,** 2199-206.

DEGNER, S. C., VERMA-GAUR, J., WONG, T. P., BOSSEN, C., IVERSON, G. M., TORKAMANI, A., VETTERMANN, C., LIN, Y. C., JU, Z., SCHULZ, D., MURRE, C. S., BIRSHTEIN, B. K., SCHORK, N. J., SCHLISSEL, M. S., RIBLET, R., MURRE, C. & FEENEY, A. J. 2011. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci U S A,* 108**,** 9566-71.

DI NOIA, J. M. & NEUBERGER, M. S. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem,* 76**,** 1-22.

DICKERSON, S. K., MARKET, E., BESMER, E. & PAPAVASILIOU, F. N. 2003. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med,* 197**,** 1291-6.

DIEFENBACH, A., JENSEN, E. R., JAMIESON, A. M. & RAULET, D. H. 2001. Rae1 and H60 ligands of the NKG2D receptor stimulate tumour immunity. *Nature,* 413**,** 165-71.

DORNER, T., FOSTER, S. J., FARNER, N. L. & LIPSKY, P. E. 1998. Somatic hypermutation of human immunoglobulin heavy chain genes: targeting of RGYW motifs on both DNA strands. *Eur J Immunol,* 28**,** 3384-96.

DORSETT, Y., MCBRIDE, K. M., JANKOVIC, M., GAZUMYAN, A., THAI, T. H., ROBBIANI, D. F., DI VIRGILIO, M., REINA SAN-MARTIN, B., HEIDKAMP, G., SCHWICKERT, T. A., EISENREICH, T., RAJEWSKY, K. & NUSSENZWEIG, M. C. 2008. MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity,* 28**,** 630-8.

EL MARJOU, F., JANSSEN, K. P., CHANG, B. H., LI, M., HINDIE, V., CHAN, L., LOUVARD, D., CHAMBON, P., METZGER, D. & ROBINE, S. 2004. Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis,* 39**,** 186-93.

ELGUETA, R., BENSON, M. J., DE VRIES, V. C., WASIUK, A., GUO, Y. & NOELLE, R. J. 2009. Molecular mechanism and function of CD40/CD40L engagement in the immune system. *Immunol Rev,* 229**,** 152-72.

ENDO, Y., MARUSAWA, H., KINOSHITA, K., MORISAWA, T., SAKURAI, T., OKAZAKI, I. M., WATASHI, K., SHIMOTOHNO, K., HONJO, T. & CHIBA, T. 2007. Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene,* 26**,** 5587-95.

ENDO, Y., MARUSAWA, H., KOU, T., NAKASE, H., FUJII, S., FUJIMORI, T., KINOSHITA, K., HONJO, T. & CHIBA, T. 2008. Activation-induced cytidine deaminase links between inflammation and the development of colitis-associated colorectal cancers. *Gastroenterology,* 135**,** 889-98, 898 e1-3.

FEAGINS, L. A., SOUZA, R. F. & SPECHLER, S. J. 2009. Carcinogenesis in IBD: potential targets for the prevention of colorectal cancer. *Nat Rev Gastroenterol Hepatol,* 6**,** 297-305.

FEDORIW, A. M., STEIN, P., SVOBODA, P., SCHULTZ, R. M. & BARTOLOMEI, M. S. 2004. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science,* 303**,** 238-40.

FILIPPOVA, G. N., FAGERLIE, S., KLENOVA, E. M., MYERS, C., DEHNER, Y., GOODWIN, G., NEIMAN, P. E., COLLINS, S. J. & LOBANENKOV, V. V. 1996. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol Cell Biol,* 16**,** 2802-13.

FODDE, R. & SMITS, R. 2001. Disease model: familial adenomatous polyposis. *Trends Mol Med,* 7**,** 369-73.

FUKITA, Y., JACOBS, H. & RAJEWSKY, K. 1998. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity,* 9**,** 105-14.

GASSER, S., ORSULIC, S., BROWN, E. J. & RAULET, D. H. 2005. The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. *Nature,* 436**,** 1186-90.

GASSER, S. & RAULET, D. 2006. The DNA damage response, immunity and cancer. *Semin Cancer Biol,* 16**,** 344-7.

GERASIMOVA, T., GUO, C., GHOSH, A., QIU, X., MONTEFIORI, L., VERMA-GAUR, J., CHOI, N. M., FEENEY, A. J. & SEN, R. 2015. A structural hierarchy mediated by multiple nuclear factors establishes IgH locus conformation. *Genes Dev,* 29**,** 1683-95.

GILES, K. E., GOWHER, H., GHIRLANDO, R., JIN, C. & FELSENFELD, G. 2010. Chromatin boundaries, insulators, and long-range interactions in the nucleus. *Cold Spring Harb Symp Quant Biol,* 75**,** 79-85.

GOLAN-MASHIACH, M., GRUNSPAN, M., EMMANUEL, R., GIBBS-BAR, L., DIKSTEIN, R. & SHAPIRO, E. 2012. Identification of CTCF as a master regulator of the clustered protocadherin genes. *Nucleic Acids Res,* 40**,** 3378-91.

GOURZI, P., LEONOVA, T. & PAPAVASILIOU, F. N. 2006. A role for activation-induced cytidine deaminase in the host response against a transforming retrovirus. *Immunity,* 24**,** 779-86.

GOUT, J., POMMIER, R. M., VINCENT, D. F., KANIEWSKI, B., MARTEL, S., VALCOURT, U. & BARTHOLIN, L. 2013. Isolation and culture of mouse primary pancreatic acinar cells. *J Vis Exp*.

GUERRA, N., TAN, Y. X., JONCKER, N. T., CHOY, A., GALLARDO, F., XIONG, N., KNOBLAUGH, S., CADO, D., GREENBERG, N. M. & RAULET, D. H. 2008. NKG2D-deficient mice are

defective in tumor surveillance in models of spontaneous malignancy. *Immunity,* 28**,** 571-80.

GUO, C., YOON, H. S., FRANKLIN, A., JAIN, S., EBERT, A., CHENG, H. L., HANSEN, E., DESPO, O., BOSSEN, C., VETTERMANN, C., BATES, J. G., RICHARDS, N., MYERS, D., PATEL, H., GALLAGHER, M., SCHLISSEL, M. S., MURRE, C., BUSSLINGER, M., GIALLOURAKIS, C. C. & ALT, F. W. 2011. CTCF-binding elements mediate control of V(D)J recombination. *Nature,* 477**,** 424-30.

GUO, Y., MONAHAN, K., WU, H., GERTZ, J., VARLEY, K. E., LI, W., MYERS, R. M., MANIATIS, T. & WU, Q. 2012. CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. *Proc Natl Acad Sci U S A,* 109**,** 21081-6.

HANDOKO, L., XU, H., LI, G., NGAN, C. Y., CHEW, E., SCHNAPP, M., LEE, C. W., YE, C., PING, J. L., MULAWADI, F., WONG, E., SHENG, J., ZHANG, Y., POH, T., CHAN, C. S., KUNARSO, G., SHAHAB, A., BOURQUE, G., CACHEUX-RATABOUL, V., SUNG, W. K., RUAN, Y. & WEI, C. L. 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet,* 43**,** 630-8.

HAUSER, A. E., JUNT, T., MEMPEL, T. R., SNEDDON, M. W., KLEINSTEIN, S. H., HENRICKSON, S. E., VON ANDRIAN, U. H., SHLOMCHIK, M. J. & HABERMAN, A. M. 2007. Definition of germinal-center B cell migration in vivo reveals predominant intrazonal circulation patterns. *Immunity,* 26**,** 655-67.

HEATH, H., RIBEIRO DE ALMEIDA, C., SLEUTELS, F., DINGJAN, G., VAN DE NOBELEN, S., JONKERS, I., LING, K. W., GRIBNAU, J., RENKAWITZ, R., GROSVELD, F., HENDRIKS, R. W. & GALJART, N. 2008. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *EMBO J,* 27**,** 2839-50.

HEINTZMAN, N. D., HON, G. C., HAWKINS, R. D., KHERADPOUR, P., STARK, A., HARP, L. F., YE, Z., LEE, L. K., STUART, R. K., CHING, C. W., CHING, K. A., ANTOSIEWICZ-BOURGET, J. E., LIU, H., ZHANG, X., GREEN, R. D., LOBANENKOV, V. V., STEWART, R., THOMSON, J. A., CRAWFORD, G. E., KELLIS, M. & REN, B. 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature,* 459**,** 108-12.

HIRAYAMA, T., TARUSAWA, E., YOSHIMURA, Y., GALJART, N. & YAGI, T. 2012. CTCF is required for neural development and stochastic expression of clustered Pcdh genes in neurons. *Cell Rep,* 2**,** 345-57.

ITO, S., NAGAOKA, H., SHINKURA, R., BEGUM, N., MURAMATSU, M., NAKATA, M. & HONJO, T. 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci U S A,* 101**,** 1975-80.

JOHNSON-LEGER, C., CHRISTENSEN, J. & KLAUS, G. G. 1998. CD28 co-stimulation stabilizes the expression of the CD40 ligand on T cells. *Int Immunol,* 10**,** 1083-91.

JUNG, S., RAJEWSKY, K. & RADBRUCH, A. 1993. Shutdown of class switch recombination by deletion of a switch region control element. *Science,* 259**,** 984-7.

KAWAGUCHI, Y., COOPER, B., GANNON, M., RAY, M., MACDONALD, R. J. & WRIGHT, C. V. 2002. The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet,* 32**,** 128-34.

KEHAYOVA, P., MONAHAN, K., CHEN, W. & MANIATIS, T. 2011. Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. *Proc Natl Acad Sci U S A,* 108**,** 17195-200.

KELLUM, R. & SCHEDL, P. 1991. A position-effect assay for boundaries of higher order chromosomal domains. *Cell,* 64**,** 941-50.

KIM, T. H., ABDULLAEV, Z. K., SMITH, A. D., CHING, K. A., LOUKINOV, D. I., GREEN, R. D., ZHANG, M. Q., LOBANENKOV, V. V. & REN, B. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell,* 128**,** 1231-45.

KLAUS, G. G., HOLMAN, M., JOHNSON-LEGER, C., CHRISTENSON, J. R. & KEHRY, M. R. 1999. Interaction of B cells with activated T cells reduces the threshold for CD40-mediated B cell activation. *Int Immunol,* 11**,** 71-9.

KLAUS, S. J., PINCHUK, L. M., OCHS, H. D., LAW, C. L., FANSLOW, W. C., ARMITAGE, R. J. & CLARK, E. A. 1994. Costimulation through CD28 enhances T cell-dependent B cell activation via CD40-CD40L interaction. *J Immunol,* 152**,** 5643-52.

KLENOVA, E. M., NICOLAS, R. H., PATERSON, H. F., CARNE, A. F., HEATH, C. M., GOODWIN, G. H., NEIMAN, P. E. & LOBANENKOV, V. V. 1993. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol Cell Biol,* 13**,** 7612-24.

KNODEL, M., KUSS, A. W., BERBERICH, I. & SCHIMPL, A. 2001. Blimp-1 over-expression abrogates IL-4- and CD40-mediated suppression of terminal B cell differentiation but arrests isotype switching. *Eur J Immunol,* 31**,** 1972-80.

KOMORI, J., MARUSAWA, H., MACHIMOTO, T., ENDO, Y., KINOSHITA, K., KOU, T., HAGA, H., IKAI, I., UEMOTO, S. & CHIBA, T. 2008. Activation-induced cytidine deaminase links bile duct inflammation to human cholangiocarcinoma. *Hepatology,* 47**,** 888-96.

KOU, T., MARUSAWA, H., KINOSHITA, K., ENDO, Y., OKAZAKI, I. M., UEDA, Y., KODAMA, Y., HAGA, H., IKAI, I. & CHIBA, T. 2007. Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int J Cancer,* 120**,** 469-76.

KOVALCHUK, A. L., DUBOIS, W., MUSHINSKI, E., MCNEIL, N. E., HIRT, C., QI, C. F., LI, Z., JANZ, S., HONJO, T., MURAMATSU, M., RIED, T., BEHRENS, T. & POTTER, M. 2007. AID-deficient Bcl-xL transgenic mice develop delayed atypical plasma cell tumors with unusual Ig/Myc chromosomal rearrangements. *J Exp Med,* 204**,** 2989-3001.

KRUEGER, F., ANDREWS, S. R. & OSBORNE, C. S. 2011. Large scale loss of data in low-diversity illumina sequencing libraries can be recovered by deferred cluster calling. *PLoS One,* 6**,** e16607.

KUNARSO, G., CHIA, N. Y., JEYAKANI, J., HWANG, C., LU, X., CHAN, Y. S., NG, H. H. & BOURQUE, G. 2010. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat Genet,* 42**,** 631-4.

KUPPERS, R. 2005. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer,* 5**,** 251-62.

KWON, K., HUTTER, C., SUN, Q., BILIC, I., COBALEDA, C., MALIN, S. & BUSSLINGER, M. 2008. Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. *Immunity,* 28**,** 751-62.

LI, B. & DEWEY, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics,* 12**,** 323.

LIU, M., DUKE, J. L., RICHTER, D. J., VINUESA, C. G., GOODNOW, C. C., KLEINSTEIN, S. H. & SCHATZ, D. G. 2008. Two levels of protection for the B cell genome during somatic hypermutation. *Nature,* 451**,** 841-5.

LOBANENKOV, V. V., NICOLAS, R. H., ADLER, V. V., PATERSON, H., KLENOVA, E. M., POLOTSKAJA, A. V. & GOODWIN, G. H. 1990. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene,* 5**,** 1743-53.

MACDONALD, R. J., SWIFT, G. H. & REAL, F. X. 2010. Transcriptional control of acinar development and homeostasis. *Prog Mol Biol Transl Sci,* 97**,** 1-40.

MACLENNAN, I. C. 1994. Germinal centers. *Annu Rev Immunol,* 12**,** 117-39.

MAJUMDER, P. & BOSS, J. M. 2010. CTCF controls expression and chromatin architecture of the human major histocompatibility complex class II locus. *Mol Cell Biol,* 30**,** 4211-23.

MAJUMDER, P., GOMEZ, J. A., CHADWICK, B. P. & BOSS, J. M. 2008. The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. *J Exp Med,* 205**,** 785-98.

MANSER, T. 1987. Mitogen-driven B cell proliferation and differentiation are not accompanied by hypermutation of immunoglobulin variable region genes. *J Immunol,* 139**,** 234-8.

MARKIEWICZ, M. A., WISE, E. L., BUCHWALD, Z. S., PINTO, A. K., ZAFIROVA, B., POLIC, B. & SHAW, A. S. 2012. RAE1epsilon ligand expressed on pancreatic islets recruits NKG2D receptor-expressing cytotoxic T cells independent of T cell receptor recognition. *Immunity,* 36**,** 132-41.

MARTINELLI, P., MADRILES, F., CANAMERO, M., PAU, E. C., POZO, N. D., GUERRA, C. & REAL, F. X. 2016. The acinar regulator Gata6 suppresses KrasG12V-driven pancreatic tumorigenesis in mice. *Gut,* 65**,** 476-86.

MARUSAWA, H., TAKAI, A. & CHIBA, T. 2011. Role of activation-induced cytidine deaminase in inflammation-associated cancer development. *Adv Immunol,* 111**,** 109-41.

MATSUMOTO, Y., MARUSAWA, H., KINOSHITA, K., ENDO, Y., KOU, T., MORISAWA, T., AZUMA, T., OKAZAKI, I. M., HONJO, T. & CHIBA, T. 2007. Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med,* 13**,** 470-6.

MAUL, R. W., SARIBASAK, H., MARTOMO, S. A., MCCLURE, R. L., YANG, W., VAISMAN, A., GRAMLICH, H. S., SCHATZ, D. G., WOODGATE, R., WILSON, D. M., 3RD & GEARHART, P. J. 2011. Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions. *Nat Immunol,* 12**,** 70-6.

MCBRIDE, K. M., BARRETO, V., RAMIRO, A. R., STAVROPOULOS, P. & NUSSENZWEIG, M. C. 2004. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J Exp Med,* 199**,** 1235-44.

MCBRIDE, K. M., GAZUMYAN, A., WOO, E. M., SCHWICKERT, T. A., CHAIT, B. T. & NUSSENZWEIG, M. C. 2008. Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J Exp Med,* 205**,** 2585-94.

MEDVEDOVIC, J., EBERT, A., TAGOH, H., TAMIR, I. M., SCHWICKERT, T. A., NOVATCHKOVA, M., SUN, Q., HUIS IN 'T VELD, P. J., GUO, C., YOON, H. S., DENIZOT, Y., HOLWERDA, S. J., DE LAAT, W., COGNE, M., SHI, Y., ALT, F. W. & BUSSLINGER, M. 2013. Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. *Immunity,* 39**,** 229-44.

MENG, F. L., DU, Z., FEDERATION, A., HU, J., WANG, Q., KIEFFER-KWON, K. R., MEYERS, R. M., AMOR, C., WASSERMAN, C. R., NEUBERG, D., CASELLAS, R., NUSSENZWEIG, M. C.,

BRADNER, J. E., LIU, X. S. & ALT, F. W. 2014. Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell,* 159**,** 1538-48.

METHOT, S. P., LITZLER, L. C., TRAJTENBERG, F., ZAHN, A., ROBERT, F., PELLETIER, J., BUSCHIAZZO, A., MAGOR, B. G. & DI NOIA, J. M. 2015. Consecutive interactions with HSP90 and eEF1A underlie a functional maturation and storage pathway of AID in the cytoplasm. *J Exp Med,* 212**,** 581-96.

MONAHAN, K., RUDNICK, N. D., KEHAYOVA, P. D., PAULI, F., NEWBERRY, K. M., MYERS, R. M. & MANIATIS, T. 2012. Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin-alpha gene expression. *Proc Natl Acad Sci U S A,* 109**,** 9125-30.

MURAMATSU, M., KINOSHITA, K., FAGARASAN, S., YAMADA, S., SHINKAI, Y. & HONJO, T. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell,* 102**,** 553-63.

MURAMATSU, M., SANKARANAND, V. S., ANANT, S., SUGAI, M., KINOSHITA, K., DAVIDSON, N. O. & HONJO, T. 1999. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem,* 274**,** 18470-6.

MUTO, T., OKAZAKI, I. M., YAMADA, S., TANAKA, Y., KINOSHITA, K., MURAMATSU, M., NAGAOKA, H. & HONJO, T. 2006. Negative regulation of activation-induced cytidine deaminase in B cells. *Proc Natl Acad Sci U S A,* 103**,** 2752-7.

NYABI, O., NAESSENS, M., HAIGH, K., GEMBARSKA, A., GOOSSENS, S., MAETENS, M., DE CLERCQ, S., DROGAT, B., HAENEBALCKE, L., BARTUNKOVA, S., DE VOS, I., DE CRAENE, B., KARIMI, M., BERX, G., NAGY, A., HILSON, P., MARINE, J. C. & HAIGH, J. J. 2009. Efficient mouse transgenesis using Gateway-compatible ROSA26 locus targeting vectors and F1 hybrid ES cells. *Nucleic Acids Res,* 37**,** e55.

OKAZAKI, I. M., HIAI, H., KAKAZU, N., YAMADA, S., MURAMATSU, M., KINOSHITA, K. & HONJO, T. 2003. Constitutive expression of AID leads to tumorigenesis. *J Exp Med,* 197**,** 1173-81.

ONG, C. T. & CORCES, V. G. 2014. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet,* 15**,** 234-46.

PASQUALUCCI, L., BHAGAT, G., JANKOVIC, M., COMPAGNO, M., SMITH, P., MURAMATSU, M., HONJO, T., MORSE, H. C., 3RD, NUSSENZWEIG, M. C. & DALLA-FAVERA, R. 2008. AID is required for germinal center-derived lymphomagenesis. *Nat Genet,* 40**,** 108-12.

PASQUALUCCI, L., NEUMEISTER, P., GOOSSENS, T., NANJANGUD, G., CHAGANTI, R. S., KUPPERS, R. & DALLA-FAVERA, R. 2001. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature,* 412**,** 341-6.

PAVRI, R., GAZUMYAN, A., JANKOVIC, M., DI VIRGILIO, M., KLEIN, I., ANSARAH-SOBRINHO, C., RESCH, W., YAMANE, A., REINA SAN-MARTIN, B., BARRETO, V., NIELAND, T. J., ROOT, D. E., CASELLAS, R. & NUSSENZWEIG, M. C. 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell,* 143**,** 122-33.

PAVRI, R. & NUSSENZWEIG, M. C. 2011. AID targeting in antibody diversity. *Adv Immunol,* 110**,** 1-26.

PEFANIS, E., WANG, J., ROTHSCHILD, G., LIM, J., CHAO, J., RABADAN, R., ECONOMIDES, A. N. & BASU, U. 2014. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature,* 514**,** 389-93.

PEREZ-DURAN, P., BELVER, L., DE YEBENES, V. G., DELGADO, P., PISANO, D. G. & RAMIRO, A. R. 2012. UNG shapes the specificity of AID-induced somatic hypermutation. *J Exp Med,* 209**,** 1379-89.

PETERS, A. & STORB, U. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity,* 4**,** 57-65.

PETERSEN-MAHRT, S. K., HARRIS, R. S. & NEUBERGER, M. S. 2002. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature,* 418**,** 99-103.

PHAM, P., BRANSTEITTER, R., PETRUSKA, J. & GOODMAN, M. F. 2003. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature,* 424**,** 103-7.

PHILLIPS, J. E. & CORCES, V. G. 2009. CTCF: master weaver of the genome. *Cell,* 137**,** 1194-211.

PISKURICH, J. F., LIN, K. I., LIN, Y., WANG, Y., TING, J. P. & CALAME, K. 2000. BLIMP-I mediates extinction of major histocompatibility class II transactivator expression in plasma cells. *Nat Immunol,* 1**,** 526-32.

RADA, C. & MILSTEIN, C. 2001. The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially. *EMBO J,* 20**,** 4570-6.

RAMIRO, A. R., JANKOVIC, M., CALLEN, E., DIFILIPPANTONIO, S., CHEN, H. T., MCBRIDE, K. M., EISENREICH, T. R., CHEN, J., DICKINS, R. A., LOWE, S. W., NUSSENZWEIG, A. & NUSSENZWEIG, M. C. 2006. Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. *Nature,* 440**,** 105-9.

RAMIRO, A. R., JANKOVIC, M., EISENREICH, T., DIFILIPPANTONIO, S., CHEN-KIANG, S., MURAMATSU, M., HONJO, T., NUSSENZWEIG, A. & NUSSENZWEIG, M. C. 2004. AID is required for c-myc/IgH chromosome translocations in vivo. *Cell,* 118**,** 431-8.

RAMIRO, A. R., STAVROPOULOS, P., JANKOVIC, M. & NUSSENZWEIG, M. C. 2003. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat Immunol,* 4**,** 452-6.

RANDALL, T. D., HEATH, A. W., SANTOS-ARGUMEDO, L., HOWARD, M. C., WEISSMAN, I. L. & LUND, F. E. 1998. Arrest of B lymphocyte terminal differentiation by CD40 signaling: mechanism for lack of antibody-secreting cells in germinal centers. *Immunity,* 8**,** 733-42.

RAULET, D. H., GASSER, S., GOWEN, B. G., DENG, W. & JUNG, H. 2013. Regulation of ligands for the NKG2D activating receptor. *Annu Rev Immunol,* 31**,** 413-41.

REVY, P., MUTO, T., LEVY, Y., GEISSMANN, F., PLEBANI, A., SANAL, O., CATALAN, N., FORVEILLE, M., DUFOURCQ-LABELOUSE, R., GENNERY, A., TEZCAN, I., ERSOY, F., KAYSERILI, H., UGAZIO, A. G., BROUSSE, N., MURAMATSU, M., NOTARANGELO, L. D., KINOSHITA, K., HONJO, T., FISCHER, A. & DURANDY, A. 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell,* 102**,** 565-75.

RIBEIRO DE ALMEIDA, C., STADHOUDERS, R., DE BRUIJN, M. J., BERGEN, I. M., THONGJUEA, S., LENHARD, B., VAN IJCKEN, W., GROSVELD, F., GALJART, N., SOLER, E. & HENDRIKS,

R. W. 2011. The DNA-binding protein CTCF limits proximal Vkappa recombination and restricts kappa enhancer interactions to the immunoglobulin kappa light chain locus. *Immunity,* 35**,** 501-13.

ROBBIANI, D. F., BOTHMER, A., CALLEN, E., REINA-SAN-MARTIN, B., DORSETT, Y., DIFILIPPANTONIO, S., BOLLAND, D. J., CHEN, H. T., CORCORAN, A. E., NUSSENZWEIG, A. & NUSSENZWEIG, M. C. 2008. AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. *Cell,* 135**,** 1028-38.

ROBBIANI, D. F., BUNTING, S., FELDHAHN, N., BOTHMER, A., CAMPS, J., DEROUBAIX, S., MCBRIDE, K. M., KLEIN, I. A., STONE, G., EISENREICH, T. R., RIED, T., NUSSENZWEIG, A. & NUSSENZWEIG, M. C. 2009. AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. *Mol Cell,* 36**,** 631-41.

ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.

ROGOZIN, I. B. & KOLCHANOV, N. A. 1992. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim Biophys Acta,* 1171**,** 11-8.

SANYAL, A., LAJOIE, B. R., JAIN, G. & DEKKER, J. 2012. The long-range interaction landscape of gene promoters. *Nature,* 489**,** 109-13.

SCHMIDT, D., SCHWALIE, P. C., WILSON, M. D., BALLESTER, B., GONCALVES, A., KUTTER, C., BROWN, G. D., MARSHALL, A., FLICEK, P. & ODOM, D. T. 2012. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell,* 148**,** 335-48.

SCHMITZ, R., YOUNG, R. M., CERIBELLI, M., JHAVAR, S., XIAO, W., ZHANG, M., WRIGHT, G., SHAFFER, A. L., HODSON, D. J., BURAS, E., LIU, X., POWELL, J., YANG, Y., XU, W., ZHAO, H., KOHLHAMMER, H., ROSENWALD, A., KLUIN, P., MULLER-HERMELINK, H. K., OTT, G., GASCOYNE, R. D., CONNORS, J. M., RIMSZA, L. M., CAMPO, E., JAFFE, E. S., DELABIE, J., SMELAND, E. B., OGWANG, M. D., REYNOLDS, S. J., FISHER, R. I., BRAZIEL, R. M., TUBBS, R. R., COOK, J. R., WEISENBURGER, D. D., CHAN, W. C., PITTALUGA, S., WILSON, W., WALDMANN, T. A., ROWE, M., MBULAITEYE, S. M., RICKINSON, A. B. & STAUDT, L. M. 2012. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature,* 490**,** 116-20.

SERNANDEZ, I. V., DE YEBENES, V. G., DORSETT, Y. & RAMIRO, A. R. 2008. Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. *PLoS One,* 3**,** e3927.

SHAFFER, A. L., LIN, K. I., KUO, T. C., YU, X., HURT, E. M., ROSENWALD, A., GILTNANE, J. M., YANG, L., ZHAO, H., CALAME, K. & STAUDT, L. M. 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity,* 17**,** 51-62.

SHAPIRO-SHELEF, M., LIN, K. I., MCHEYZER-WILLIAMS, L. J., LIAO, J., MCHEYZER-WILLIAMS, M. G. & CALAME, K. 2003. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity,* 19**,** 607-20.

SHEN, H. M., PETERS, A., BARON, B., ZHU, X. & STORB, U. 1998. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science,* 280**,** 1750-2.

SHI, W., LIAO, Y., WILLIS, S. N., TAUBENHEIM, N., INOUYE, M., TARLINTON, D. M., SMYTH, G. K., HODGKIN, P. D., NUTT, S. L. & CORCORAN, L. M. 2015. Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. *Nat Immunol,* 16**,** 663-73.

SILACCI, P., MOTTET, A., STEIMLE, V., REITH, W. & MACH, B. 1994. Developmental extinction of major histocompatibility complex class II gene expression in plasmocytes is mediated by silencing of the transactivator gene CIITA. *J Exp Med,* 180**,** 1329-36.

STAVNEZER-NORDGREN, J. & SIRLIN, S. 1986. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. *EMBO J,* 5**,** 95-102.

STAVNEZER, J. 2011. Complex regulation and function of activation-induced cytidine deaminase. *Trends Immunol,* 32**,** 194-201.

STAVNEZER, J., GUIKEMA, J. E. & SCHRADER, C. E. 2008. Mechanism and regulation of class switch recombination. *Annu Rev Immunol,* 26**,** 261-92.

TAKAI, A., MARUSAWA, H., MINAKI, Y., WATANABE, T., NAKASE, H., KINOSHITA, K., TSUJIMOTO, G. & CHIBA, T. 2012. Targeting activation-induced cytidine deaminase prevents colon cancer development despite persistent colonic inflammation. *Oncogene,* 31**,** 1733-42.

TAKAI, A., TOYOSHIMA, T., UEMURA, M., KITAWAKI, Y., MARUSAWA, H., HIAI, H., YAMADA, S., OKAZAKI, I. M., HONJO, T., CHIBA, T. & KINOSHITA, K. 2009. A novel mouse model of hepatocarcinogenesis triggered by AID causing deleterious p53 mutations. *Oncogene,* 28**,** 469-78.

TAKIZAWA, M., TOLAROVA, H., LI, Z., DUBOIS, W., LIM, S., CALLEN, E., FRANCO, S., MOSAICO, M., FEIGENBAUM, L., ALT, F. W., NUSSENZWEIG, A., POTTER, M. & CASELLAS, R. 2008. AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J Exp Med,* 205**,** 1949-57.

TENG, G., HAKIMPOUR, P., LANDGRAF, P., RICE, A., TUSCHL, T., CASELLAS, R. & PAPAVASILIOU, F. N. 2008. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity,* 28**,** 621-9.

UCHIMURA, Y., BARTON, L. F., RADA, C. & NEUBERGER, M. S. 2011. REG-gamma associates with and modulates the abundance of nuclear activation-induced deaminase. *J Exp Med,* 208**,** 2385-91.

VICTORA, G. D. & NUSSENZWEIG, M. C. 2012. Germinal centers. *Annu Rev Immunol,* 30**,** 429-57.

VONDERHEIDE, R. H. & BAYNE, L. J. 2013. Inflammatory networks and immune surveillance of pancreatic carcinoma. *Curr Opin Immunol,* 25**,** 200-5.

WANG, H., MAURANO, M. T., QU, H., VARLEY, K. E., GERTZ, J., PAULI, F., LEE, K., CANFIELD, T., WEAVER, M., SANDSTROM, R., THURMAN, R. E., KAUL, R., MYERS, R. M. & STAMATOYANNOPOULOS, J. A. 2012. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res,* 22**,** 1680-8.

XU, Z., ZAN, H., PONE, E. J., MAI, T. & CASALI, P. 2012. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol,* 12**,** 517-31.

YAMANE, A., RESCH, W., KUO, N., KUCHEN, S., LI, Z., SUN, H. W., ROBBIANI, D. F., MCBRIDE, K., NUSSENZWEIG, M. C. & CASELLAS, R. 2011. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol,* 12**,** 62-9.

YU, K., ROY, D., BAYRAMYAN, M., HAWORTH, I. S. & LIEBER, M. R. 2005. Fine-structure analysis of activation-induced deaminase accessibility to class switch region R-loops. *Mol Cell Biol,* 25**,** 1730-6.

ANNEX

Table1. 20% more significantly downregulated genes in CTCF$^{fl/fl}$ mice compared with control mice in CD3/CD28 stimulation.

Gene	ave_CTCF $^{fl/fl}$	ave_CTCF$^{fl/+}$	foldChange	adj.P.Val
Fcer2a	4,41	34,46	-7,70	4,30E-37
Slc15a3	8,74	46,18	-5,27	1,41E-35
Zfp280b	74,49	249,15	-3,35	1,78E-33
Rnf19b	105,12	364,86	-3,47	5,12E-30
Fmnl2	3,44	23,73	-6,98	1,92E-22
Rab35	64,56	169,65	-2,63	5,00E-21
Atp11a	9,97	36,19	-3,63	8,96E-21
Lifr	11,26	46,30	-4,14	4,00E-20
Samsn1	48,64	112,00	-2,30	1,70E-17
Pros1	4,14	20,39	-4,91	4,16E-17
Insm1	23,55	91,36	-3,86	7,72E-16
Ctcf	**70,69**	**145,97**	**-2,06**	**1,93E-15**
Pkib	67,05	143,96	-2,15	4,36E-15
Tspan2	39,12	86,63	-2,21	2,89E-14
Cish	34,22	100,61	-2,94	1,12E-13
Lta	2,08	12,48	-5,91	2,18E-13
Parm1	0,46	6,95	-15,02	3,09E-13
Klhdc2	62,86	134,65	-2,14	3,09E-13
Lif	2,99	14,78	-4,88	3,59E-13
Fabp5	20,85	51,35	-2,45	5,16E-13
Batf3	5,80	25,09	-4,34	6,27E-13
Il4i1	131,62	274,27	-2,08	1,05E-12
Sema4b	46,78	109,18	-2,33	1,43E-12
Cpox	233,90	515,48	-2,20	1,43E-12
Dock10	171,23	362,34	-2,12	3,77E-12
Nfil3	68,62	153,03	-2,23	5,10E-12
Hipk2	106,07	288,25	-2,72	1,03E-11
Akap2	27,82	59,27	-2,13	1,28E-11
Mthfd2	62,69	119,41	-1,90	1,87E-11
Xylt1	2,25	12,46	-5,61	2,04E-11
Ptgs1	0,34	5,86	-16,59	2,45E-11
Thy1	14,27	49,35	-3,43	2,80E-11
Hlx	3,11	14,54	-4,64	4,15E-11
Pprc1	87,03	166,58	-1,91	5,37E-11
Marcks	27,39	61,48	-2,23	8,99E-11
Mat2a	151,33	293,96	-1,94	4,35E-10
Ifng	6,59	21,25	-3,21	4,84E-10
Wdr77	45,39	83,32	-1,83	1,09E-09

Gm26667	7,20	21,72	-2,97	1,73E-09
Sypl	50,93	95,61	-1,88	1,79E-09
Tmed8	92,78	173,71	-1,87	2,27E-09
Gnl3	64,95	122,33	-1,88	2,60E-09
Naa25	93,21	172,22	-1,85	2,69E-09
Pim3	39,97	75,01	-1,88	2,98E-09
Snhg4	17,91	39,08	-2,17	3,21E-09
Otulin	104,54	200,20	-1,91	3,51E-09
Polr1b	66,60	118,04	-1,77	6,82E-09
Gfi1	90,63	155,53	-1,72	7,78E-09
Wdr59	26,40	51,02	-1,93	1,42E-08
Nol9	87,73	151,63	-1,73	1,43E-08
Gm17300	8,01	22,25	-2,75	1,56E-08
Gzma	1,55	10,98	-7,04	1,61E-08
Abcb1b	20,07	41,78	-2,07	2,10E-08
Ass1	27,33	52,96	-1,94	2,10E-08
Utp15	56,51	97,33	-1,72	2,16E-08
Rrp12	58,65	110,12	-1,88	3,58E-08
Rbpj	69,77	128,20	-1,83	3,64E-08
Epas1	2,24	9,97	-4,47	3,69E-08
Rab8b	62,75	104,62	-1,67	3,87E-08
Egr2	52,54	98,41	-1,88	3,96E-08
Syngr2	336,51	624,75	-1,86	4,06E-08
Eef1e1	33,18	59,70	-1,80	4,27E-08
Ak2	170,51	319,03	-1,87	5,11E-08
Ptpn3	7,75	20,07	-2,60	6,20E-08
St7	12,12	27,68	-2,28	6,34E-08
Ing5	23,78	46,39	-1,95	7,39E-08
Asns	32,02	60,43	-1,88	1,00E-07
Cd44	158,75	287,29	-1,81	1,02E-07
Iigp1	131,55	239,58	-1,82	1,12E-07
Socs1	104,86	207,08	-1,97	1,18E-07
Utp20	78,41	134,62	-1,72	1,49E-07
Ighg1	191,39	343,48	-1,79	1,57E-07
Mettl16	52,72	88,85	-1,69	1,87E-07
1110038B12Rik	23,47	46,23	-1,96	1,88E-07
St6galnac4	80,91	132,28	-1,63	2,08E-07
Jade2	72,21	123,36	-1,71	2,49E-07
Rbpsuh-rs3	9,98	22,62	-2,27	3,34E-07
Il4ra	392,87	692,40	-1,76	3,34E-07
Serpinb9	36,78	73,55	-2,00	3,89E-07
Heatr1	130,69	222,92	-1,71	4,80E-07
Wdr18	53,22	90,11	-1,69	4,88E-07

Ccl22	28,91	57,79	-1,99	5,10E-07
Nabp1	94,36	181,78	-1,93	5,22E-07
1110032F04Rik	5,29	14,38	-2,69	5,80E-07
Slc16a1	89,59	144,50	-1,61	7,17E-07
Slc39a6	29,91	53,51	-1,79	8,07E-07
Srm	124,28	227,15	-1,83	8,56E-07
Grwd1	38,16	68,36	-1,79	8,91E-07
Cirh1a	70,85	113,99	-1,61	9,09E-07
Ado	35,81	62,80	-1,76	9,58E-07
Eftud1	34,45	58,74	-1,71	9,80E-07
Atf3	5,59	14,45	-2,60	1,04E-06
Nfkb2	167,60	298,03	-1,78	1,15E-06
Pop1	31,87	57,55	-1,80	1,20E-06
Dkc1	94,12	155,42	-1,65	1,33E-06
Dusp10	11,27	25,81	-2,30	1,55E-06
Wdr75	58,90	95,25	-1,62	1,67E-06
Pno1	35,54	63,18	-1,78	1,70E-06
Mfhas1	42,95	73,06	-1,70	1,82E-06
Spred1	13,59	27,73	-2,04	2,29E-06
Slpi	3,07	10,03	-3,27	2,31E-06
Brix1	56,07	89,89	-1,60	2,51E-06
2410002F23Rik	56,14	92,65	-1,65	2,61E-06
Amd1	73,88	117,23	-1,59	2,62E-06
Heatr3	60,80	98,09	-1,61	2,73E-06
Lad1	1,03	5,62	-5,43	2,82E-06
Timd2	5,08	13,88	-2,71	2,82E-06
Zc3h12c	2,25	8,96	-4,03	3,26E-06
Mdn1	145,61	255,90	-1,76	3,72E-06
Nol11	49,19	78,81	-1,60	3,72E-06
Alpl	70,97	112,82	-1,59	3,72E-06
Tnfrsf8	44,92	99,54	-2,21	3,98E-06
Ifrd1	23,11	44,63	-1,93	4,79E-06
Arhgap5	2,77	9,04	-3,24	4,86E-06
Dusp7	11,02	23,16	-2,10	5,60E-06
Hspd1	431,84	753,39	-1,74	5,77E-06
Rcc1	97,49	159,08	-1,63	6,29E-06
Tbxas1	0,47	4,10	-8,77	6,31E-06
Mrps18b	35,16	59,37	-1,69	6,83E-06
Noc4l	99,67	155,93	-1,56	6,83E-06
Ppargc1b	14,98	30,22	-2,02	7,41E-06
Utp18	48,21	77,71	-1,61	7,61E-06
Plagl2	125,44	202,23	-1,61	7,73E-06
Xpot	101,46	156,30	-1,54	9,68E-06

Ncl	1177,19	1930,32	-1,64	9,81E-06
Ept1	42,20	67,48	-1,60	9,92E-06
Ndufaf4	15,22	30,43	-2,00	1,12E-05
Nop2	112,29	188,82	-1,68	1,19E-05
Naa15	141,34	232,09	-1,64	1,22E-05
Slco3a1	4,96	13,98	-2,81	1,25E-05
Sema6d	20,19	39,32	-1,95	1,37E-05
Urb2	54,88	84,85	-1,55	1,39E-05
Mybbp1a	430,90	710,20	-1,65	1,48E-05
Rif1	82,81	126,27	-1,53	1,69E-05
Chchd4	23,64	42,86	-1,80	1,81E-05
Yars	101,54	158,17	-1,56	1,81E-05
2700038G22Rik	11,62	23,77	-2,05	1,83E-05
Il13	0,08	1,90	-24,37	2,01E-05
Slc7a5	199,97	350,45	-1,75	2,08E-05
Htr7	0,13	2,55	-17,26	2,14E-05
Rrp9	55,81	90,71	-1,62	2,28E-05
Wdr43	132,11	209,77	-1,59	2,31E-05
Hist2h2aa1	1,60	9,39	-5,60	2,61E-05
Slc7a1	136,74	227,51	-1,66	2,64E-05
Prps1	54,51	83,90	-1,54	2,85E-05
Grap	81,17	136,06	-1,67	2,88E-05
Slc25a32	15,73	29,64	-1,88	2,90E-05
Ipo4	58,08	92,32	-1,59	3,00E-05
St8sia6	5,21	13,26	-2,57	3,27E-05
Ccdc86	40,31	64,15	-1,59	3,78E-05
Mars2	22,84	38,72	-1,70	3,80E-05
Lipg	1,99	7,07	-3,54	3,83E-05
Ifrd2	61,36	100,20	-1,63	4,07E-05
Rrs1	67,88	108,10	-1,59	4,11E-05
Dnttip2	49,22	76,14	-1,55	4,32E-05
Abcc1	50,64	76,95	-1,52	4,32E-05
Bend3	24,17	43,67	-1,80	4,32E-05
Mrto4	62,52	99,42	-1,59	4,50E-05
Bysl	43,57	69,95	-1,60	4,55E-05
Kdm6b	85,39	136,10	-1,60	4,77E-05
Nol10	40,40	65,65	-1,62	4,79E-05
Foxn2	31,08	51,69	-1,67	4,87E-05
Sipa1l3	49,12	80,05	-1,63	5,01E-05
Akap1	24,02	43,49	-1,81	5,01E-05
Umps	75,87	115,69	-1,53	5,13E-05
Dhx33	47,96	75,13	-1,56	5,25E-05
Ddx21	320,50	516,22	-1,61	5,26E-05

Cdk5r1	13,95	27,32	-1,97	5,28E-05
Cse1l	233,37	370,67	-1,59	5,49E-05
BC035044	34,92	55,44	-1,59	5,66E-05
Stau2	7,63	16,14	-2,11	6,03E-05
Set	518,22	837,02	-1,62	6,15E-05
Tgif2	16,81	29,82	-1,78	6,20E-05
Kpna4	92,65	140,53	-1,52	6,21E-05
Bcat1	52,63	81,11	-1,54	6,30E-05
Fam136a	33,99	54,85	-1,61	6,44E-05
Tuba4a	89,52	137,86	-1,54	7,12E-05
Dnmt3a	103,36	153,67	-1,49	7,93E-05
Ski	10,42	22,05	-2,12	8,97E-05
Pag1	32,54	52,14	-1,61	9,15E-05
Cdk6	27,16	45,04	-1,66	9,62E-05
Nomo1	109,90	170,98	-1,56	9,76E-05
Slc16a6	32,76	52,49	-1,60	9,89E-05
Hist1h2bg	2,24	7,56	-3,40	1,01E-04
Uchl3	30,09	49,22	-1,63	1,02E-04
Rrp15	29,32	49,61	-1,69	1,04E-04
Havcr1	3,41	9,47	-2,79	1,05E-04
BC037034	13,78	29,57	-2,16	1,10E-04
Gadd45gip1	40,45	61,37	-1,52	1,11E-04
Lpp	37,88	72,68	-1,93	1,11E-04
Pogk	17,72	32,24	-1,82	1,12E-04
Cflar	123,22	188,83	-1,53	1,14E-04
Gar1	50,44	77,90	-1,54	1,14E-04
Ccdc71l	63,11	100,38	-1,59	1,16E-04
Cluh	139,66	224,58	-1,61	1,18E-04
Ddx18	77,95	122,31	-1,57	1,21E-04
D19Bwg1357e	73,59	110,43	-1,50	1,24E-04
Fastkd5	9,44	18,50	-1,96	1,25E-04
Gpr33	0,83	4,33	-4,96	1,25E-04
Nudt4	54,14	84,69	-1,56	1,25E-04
Tnfaip3	103,57	156,81	-1,52	1,26E-04
Tsr1	104,75	153,57	-1,47	1,26E-04

Table2. 20% more significantly upregulated genes in CTCFfl/fl mice compared with control mice in CD3/CD28 stimulation.

Gene	ave_CTCF $^{fl/fl}$	ave_CTCF$^{fl/+}$	foldChange	adj.P.Val
Ccr6	108,75	28,07	3,86	5,15E-41
Sell	153,09	43,51	3,52	1,77E-38
Spata13	92,40	25,21	3,67	9,60E-32
Flt3	15,55	1,33	11,60	1,44E-26
Pou6f1	159,16	54,77	2,91	2,04E-25
Txnip	193,52	62,54	3,10	2,04E-24
Cd24a	622,11	194,05	3,21	2,04E-24
Cd97	121,49	47,35	2,57	7,00E-22
Bst1	73,20	24,52	2,98	5,00E-21
Acss1	88,53	35,25	2,51	1,49E-20
Stac2	18,84	3,44	5,43	1,58E-18
Arl5c	95,96	40,50	2,37	3,19E-18
Cd2	115,12	52,70	2,19	1,03E-17
Gpr114	22,99	4,67	4,83	1,19E-17
Aldoc	104,71	19,83	5,30	2,23E-17
Neurl3	65,99	19,68	3,32	8,53E-17
Gramd2	45,33	12,50	3,66	2,53E-16
Ifitm1	12,46	1,24	9,85	7,72E-16
Itgam	22,38	3,83	5,81	9,74E-16
Cytip	625,36	240,10	2,60	1,07E-15
Fcgr2b	219,96	92,64	2,37	2,17E-15
Trp53inp1	132,51	56,85	2,33	3,29E-15
Dok3	103,62	43,16	2,40	2,26E-14
Cyp4f18	25,67	5,59	4,52	2,89E-14
Fcrla	199,21	83,01	2,40	4,71E-14
Nod1	33,60	10,44	3,22	4,71E-14
Arhgef3	78,62	36,54	2,15	5,35E-14
Rhoh	209,78	95,36	2,20	6,15E-14
S100a10	58,63	24,49	2,40	7,72E-14
Susd3	48,87	20,29	2,42	1,14E-13
Ypel3	69,20	29,14	2,37	1,50E-13
Nkg7	36,86	10,24	3,55	1,60E-13
Tg	65,27	27,72	2,37	1,39E-12
Sfn	58,63	26,28	2,24	1,40E-12
Slc4a8	21,37	5,12	4,11	2,07E-12
Nr4a1	107,14	48,95	2,18	2,65E-12
Bank1	143,22	70,53	2,04	5,39E-12
Trib2	24,13	7,42	3,26	7,59E-12
Myl12b	90,23	45,34	1,99	7,75E-12

Ret	6,36	0,46	13,15	9,11E-12
Hes1	20,69	3,83	5,36	9,75E-12
Bzrap1	57,82	24,72	2,33	1,29E-11
Tcp11l2	33,89	7,18	4,75	1,43E-11
Fgd3	75,20	37,68	2,00	1,77E-11
Pirb	21,01	6,74	3,12	1,04E-10
Kif19a	12,84	2,94	4,44	1,05E-10
Sidt1	25,51	8,50	3,03	1,39E-10
Rhobtb2	184,61	84,36	2,19	2,23E-10
Prdm1	60,55	18,16	3,36	3,05E-10
Haao	61,47	30,25	2,03	4,84E-10
Apold1	9,57	1,31	7,51	5,70E-10
Blk	159,08	83,13	1,91	7,12E-10
Sla	138,40	76,63	1,81	7,63E-10
Pik3ip1	27,88	9,46	2,93	7,66E-10
Vps37b	231,91	117,01	1,98	7,72E-10
Rtn4ip1	80,15	38,05	2,10	7,96E-10
Erdr1	471,48	236,24	2,00	9,20E-10
Sat1	57,52	28,73	2,00	1,09E-09
Cxxc5	36,99	15,46	2,39	1,09E-09
4632428N05Rik	38,03	14,75	2,58	1,09E-09
Anxa2	40,85	13,83	2,97	1,17E-09
Abi3	88,23	46,64	1,89	1,25E-09
Ier5l	44,93	18,89	2,36	1,44E-09
Tspan13	157,12	74,45	2,11	1,67E-09
Cd86	190,53	92,12	2,07	1,86E-09
Unc119b	272,04	145,61	1,87	2,12E-09
Cerk	249,09	129,91	1,92	2,12E-09
Padi2	47,47	21,66	2,20	2,52E-09
Tm6sf1	96,79	54,32	1,78	2,72E-09
Il1r2	46,81	15,82	2,98	2,98E-09
Dtx1	144,29	78,99	1,83	3,91E-09
Syk	1021,60	551,32	1,85	4,54E-09
Acp5	21,55	8,27	2,61	4,60E-09
Entpd1	34,09	12,99	2,63	4,60E-09
Ptpn22	111,70	59,85	1,87	4,71E-09
Sertad4	10,79	2,35	4,53	4,71E-09
Tubb2b	26,88	11,11	2,42	5,00E-09
Anxa6	164,01	84,06	1,95	5,32E-09
Neil1	77,19	40,70	1,90	5,98E-09
Gadd45a	25,25	9,97	2,54	6,21E-09
Sdc1	38,90	18,07	2,16	6,71E-09
Pde4b	108,47	61,78	1,76	7,72E-09

Myo1f	93,64	53,62	1,75	1,35E-08
Tha1	52,66	22,02	2,39	1,38E-08
Bmf	12,98	2,69	4,85	1,40E-08
Stxbp1	78,65	41,37	1,90	1,41E-08
Stard10	37,39	15,57	2,39	1,51E-08
Ubac2	129,41	72,18	1,79	1,56E-08
Tbc1d10c	107,82	63,45	1,70	1,80E-08
Cpm	20,69	7,61	2,71	1,91E-08
Map3k9	39,72	19,10	2,08	2,16E-08
Elmo2	262,95	144,65	1,82	2,17E-08
Il12a	16,11	5,29	3,04	2,19E-08
Fcrl1	68,34	36,12	1,88	2,81E-08
Inpp5k	72,75	40,72	1,78	2,99E-08
Ccr7	317,90	169,30	1,88	2,99E-08
Rbl2	105,33	56,94	1,85	3,26E-08
Il10ra	118,33	67,68	1,75	3,86E-08
C920025E04Rik	55,92	26,51	2,12	3,86E-08
Cacna1e	60,16	23,62	2,54	3,96E-08
Mxd4	41,39	13,86	3,00	4,26E-08
Ganc	36,60	17,24	2,13	4,28E-08
Lax1	223,09	120,78	1,85	4,90E-08
Adamts10	65,55	35,88	1,83	4,97E-08
Blvrb	41,02	18,99	2,16	5,30E-08
Mgst2	18,42	5,45	3,37	5,30E-08
Ackr2	13,37	4,12	3,22	6,15E-08
Aldh2	321,70	177,41	1,81	6,17E-08
Scml4	16,54	4,81	3,42	7,23E-08
Pold4	101,88	54,60	1,87	7,60E-08
Cd3g	13,53	3,46	3,84	7,89E-08
Pitpnc1	50,62	23,67	2,14	9,52E-08
A530040E14Rik	118,44	69,27	1,71	9,74E-08
Dgka	84,09	48,83	1,73	1,18E-07
Smim14	94,01	56,25	1,67	1,39E-07
Cblb	93,48	52,86	1,77	1,51E-07
Gng2	42,60	19,90	2,13	1,55E-07
Trim7	16,04	5,55	2,94	1,69E-07
Bckdha	34,40	16,51	2,09	1,87E-07
Zfp395	66,93	37,28	1,80	1,91E-07
Prf1	36,86	15,84	2,31	2,22E-07
Cd27	29,35	10,77	2,70	2,29E-07
Ndrg1	52,14	24,02	2,17	2,64E-07
Arrb1	59,62	24,98	2,39	2,89E-07
Cd244	6,03	0,74	8,27	3,14E-07

Siglecg	194,05	112,22	1,73	3,28E-07
H2afv	65,86	37,11	1,77	3,58E-07
Dennd1c	61,49	34,11	1,80	3,58E-07
Ssbp3	107,67	65,15	1,65	3,80E-07
Ube2h	149,53	86,94	1,72	4,70E-07
Sepp1	15,57	5,46	2,82	4,88E-07
Gm16026	32,78	15,16	2,16	4,90E-07
Fyn	98,14	57,60	1,70	5,00E-07
Slc2a3	108,71	47,40	2,30	5,10E-07
Pld3	58,17	30,85	1,89	5,16E-07
Itm2a	14,79	5,56	2,66	5,27E-07
Pbxip1	89,72	52,77	1,71	6,34E-07
Cables1	42,22	20,06	2,10	9,13E-07
Abca7	121,62	76,42	1,59	9,16E-07
Pyhin1	162,83	91,46	1,78	9,32E-07
B3gnt8	19,67	7,92	2,47	9,44E-07
Emid1	13,59	4,76	2,83	9,54E-07
Rab31	26,25	11,20	2,32	9,70E-07
Acsf2	108,09	65,76	1,64	9,82E-07
Acot1	5,93	0,94	6,47	1,04E-06
Ccr5	48,25	20,52	2,35	1,04E-06
Id3	17,88	6,89	2,58	1,06E-06
Rnf167	68,36	41,62	1,64	1,15E-06
Gsn	58,73	28,97	2,03	1,15E-06
Chst12	57,41	32,95	1,74	1,19E-06
Trappc6a	66,86	40,23	1,66	1,22E-06
Aim1	63,00	35,45	1,77	1,41E-06
Palm	37,41	18,91	1,98	1,50E-06
Pycard	95,90	59,38	1,62	1,66E-06
Ildr1	25,26	10,94	2,30	1,66E-06
Tnfaip8	124,61	77,41	1,61	1,70E-06
Basp1	27,28	10,99	2,49	1,93E-06
Cenpa	267,55	157,28	1,70	1,94E-06
Tmem134	112,74	71,28	1,58	2,01E-06
Ano10	20,63	9,20	2,24	2,18E-06
Mta3	110,21	67,16	1,64	2,26E-06
Fam78a	111,16	68,24	1,63	2,35E-06
Tecpr1	45,39	24,61	1,84	2,49E-06
Derl3	27,19	12,65	2,15	2,79E-06
Pear1	55,76	31,58	1,77	2,82E-06
Plcg1	64,92	38,31	1,69	3,17E-06
Rhobtb1	31,11	13,93	2,24	3,21E-06
Scimp	10,03	3,15	3,14	3,21E-06

Lbh	254,37	141,23	1,80	3,59E-06
Padi3	4,78	0,69	6,57	3,59E-06
Atp13a2	156,81	100,16	1,57	3,83E-06
Unc13a	24,38	10,80	2,26	3,85E-06
Fgl2	156,47	94,67	1,65	4,08E-06
Ctla4	98,64	53,54	1,84	4,08E-06
Bcl2l11	502,57	279,61	1,80	4,27E-06
Slc25a23	19,37	8,07	2,39	4,29E-06
Hepacam2	14,29	5,13	2,78	4,82E-06
Tifa	23,64	11,27	2,09	4,89E-06
Triobp	121,13	72,66	1,67	4,92E-06
Hexb	47,82	27,56	1,74	4,92E-06
Csf1	215,14	124,93	1,72	5,71E-06
C130026I21Rik	97,86	61,23	1,59	5,72E-06
Hmha1	522,13	314,84	1,66	5,92E-06
Plgrkt	40,43	22,68	1,78	5,92E-06
Sipa1l1	139,32	87,67	1,59	5,96E-06
Klhl6	84,18	52,03	1,62	5,96E-06
Utrn	72,10	38,91	1,85	6,29E-06
Pde4a	28,72	13,47	2,13	6,32E-06
Bfsp2	12,61	4,42	2,86	6,32E-06
Traf3ip3	130,95	84,86	1,54	6,32E-06
Mgrn1	119,22	76,62	1,56	6,68E-06
Srgap3	31,51	13,25	2,38	6,86E-06
S100a11	55,10	30,33	1,83	7,00E-06
Gramd1a	156,32	98,01	1,59	7,17E-06
Gmfg	87,79	55,10	1,60	7,20E-06
Anxa1	3,32	0,28	10,96	7,51E-06
Gm2a	191,96	114,93	1,67	7,59E-06
Ankrd13d	23,42	11,43	2,05	7,59E-06
Pecam1	72,81	40,95	1,78	7,69E-06
Grik5	18,59	8,11	2,30	7,74E-06
Gapdhs	12,47	3,83	3,28	8,14E-06
Rgs2	12,16	4,28	2,83	8,35E-06
Ckb	23,55	8,38	2,83	8,90E-06
Arhgap25	150,87	96,73	1,56	9,33E-06
Chrna9	4,58	0,71	6,65	9,88E-06
Gm8369	47,98	25,70	1,87	1,01E-05
Ccdc102a	12,87	4,98	2,56	1,04E-05
Ifngr1	42,68	23,50	1,81	1,22E-05
Il16	153,42	98,99	1,55	1,25E-05
Hivep2	98,65	61,08	1,61	1,48E-05
Txndc16	112,29	71,43	1,57	1,51E-05

Serinc5	84,32	53,55	1,57	1,52E-05
Nxpe3	117,96	73,93	1,59	1,60E-05
Lag3	150,26	87,14	1,72	1,67E-05
Bcl7a	146,23	96,34	1,52	1,72E-05
Carns1	15,82	6,63	2,41	1,73E-05
Gm9845	6,35	1,39	4,45	1,74E-05
Ppcdc	47,73	26,45	1,81	1,83E-05
Dos	20,14	9,67	2,08	1,99E-05
Fam65c	3,62	0,41	9,20	2,04E-05
Itga3	7,40	1,91	3,82	2,04E-05
Emb	26,72	12,35	2,17	2,09E-05
Nicn1	12,23	4,72	2,57	2,09E-05
Add1	244,56	155,00	1,58	2,34E-05
Abcd1	66,36	42,12	1,58	2,47E-05
Grina	89,70	55,49	1,61	2,50E-05
Adck3	21,06	9,68	2,16	2,54E-05
Tmem71	10,96	4,21	2,58	2,55E-05
Ighg3	165,14	101,76	1,62	2,56E-05
Coro1b	289,42	180,35	1,60	2,58E-05
Mical1	131,56	87,26	1,51	2,60E-05
Igha	32,22	14,73	2,20	2,69E-05

Table3. Downregulated genes in CTCF$^{fl/fl}$ mice compared with control mice in LPS/IL4 stimulation.

Gene	ave_CTCF $^{fl/fl}$	ave_CTCF$^{fl/+}$	foldChange	adj.P.Val
Rsc1a1	0,22	2,85	-13,41	5,18E-05
Rps13-ps2	0,84	8,93	-10,56	3,52E-03
Iigp1	1,50	10,62	-7,06	8,24E-07
Serpina3f	13,49	41,71	-3,09	4,73E-10
Tgtp1	12,67	32,85	-2,59	5,74E-06
4930565N06Rik	1,25	3,16	-2,51	2,95E-02
Rn7sk	24,00	46,04	-1,92	8,09E-06
BC037034	14,69	27,75	-1,89	1,27E-02
Gm4759	23,22	42,91	-1,85	7,88E-04
Rltpr	54,57	99,99	-1,83	1,46E-20
Ctcf	67,20	120,97	-1,80	8,12E-19
Ly6c2	3,96	7,12	-1,80	3,04E-02
Malat1	350,87	550,04	-1,57	2,30E-02
Gm17132	16,74	25,98	-1,55	2,95E-02
Serpina3g	98,83	152,53	-1,54	8,09E-06
Gm15662	205,23	312,38	-1,52	7,88E-04
Gbp6	63,82	96,21	-1,51	3,71E-03
Ifi44	11,19	16,81	-1,50	1,88E-02
Gm17131	55,26	81,59	-1,48	3,52E-03
Gm26905	127,04	185,58	-1,46	1,27E-02
Papln	39,46	57,28	-1,45	3,22E-02
Cacna1e	68,96	100,01	-1,45	4,33E-03
Gbp2	450,07	652,04	-1,45	2,97E-02
Tgtp2	46,68	65,70	-1,41	3,22E-02
Cdh17	38,54	53,64	-1,39	8,11E-03
Igtp	58,71	81,56	-1,39	3,52E-03
Cytip	293,99	405,72	-1,38	1,48E-02
Gbp7	107,09	147,15	-1,37	5,13E-05
Gm26917	197,33	265,78	-1,35	6,11E-03
Zbp1	97,46	130,56	-1,34	1,11E-02
Eng	33,64	44,62	-1,33	1,90E-02
Spata13	58,82	76,86	-1,31	3,52E-03
Gbp5	122,27	159,70	-1,31	5,55E-03
Nlrc5	223,44	288,46	-1,29	3,51E-02
Cd86	92,77	116,89	-1,26	3,50E-02
Ubald2	146,06	183,74	-1,26	1,81E-02

Table3. Uregulated genes in CTCF$^{fl/fl}$ mice compared with control mice in LPS/IL4 stimulation.

Gene	ave_CTCF $^{fl/fl}$	ave_CTCF$^{fl/+}$	foldChange	adj.P.Val
Spock2	17,67	10,02	1,76	5,74E-06
Atf5	58,18	39,40	1,48	1,24E-05
Eps8l1	11,54	0,82	14,01	3,03E-05
Rpia	91,61	67,13	1,36	4,87E-05
Hist2h2aa1	2,85	0,55	5,20	5,95E-04
Fah	47,39	35,33	1,34	3,52E-03
Chchd10	52,23	40,65	1,29	1,29E-02
Egr2	43,59	32,49	1,34	1,48E-02
Vpreb3	40,66	31,87	1,28	2,95E-02
Srgap3	8,53	5,13	1,66	3,22E-02
Cpox	130,14	98,51	1,32	5,00E-02

PUBLICATIONS

EMBO Mol Med. 2015 Aug 17;7(10):1327-36. doi:10.15252/emmm.201505348.

AID-expressing epithelium is protected from oncogenic transformation by an NKG2D surveillance pathway.

Pérez-García A, Pérez-Durán P, Wossning T, Sernandez IV, Mur SM, Cañamero M, Real FX, Ramiro AR.